

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 21, 2005, 03:55:35 ; Search time 15968 Seconds  
(without alignments)  
11512.969 Million cell updates/sec

Title: US-10-608-863-1  
Perfect score: 3794  
Sequence: 1 ataaatgacgtagccgagaga.....gtgtgagtactaggagaagt 3794

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl:\*  
1: gb\_ba.\*  
2: gb\_hhg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sv.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3722	98.1	3802	9 D78579	D78579 Homo sapien
2	3653.6	96.3	3762	6 E12854	E12854 cDNA encodi
3	2875	75.8	4977	6 AX329720	AX329720 Sequence
4	2875	75.8	4977	9 HSU12767	U12767 Human mitog
5	2689.6	70.9	2714	6 AR380383	AR380383 Sequence
6	2689.6	70.9	2714	9 S81243	S81243 CHN-steroid
7	2500	65.9	4400	6 E10601	E10601 cDNA encodi
8	2500	65.9	4400	6 AX587468	AX587468 Sequence
9	2500	65.9	4400	10 RATNOR1	D38530 Rattus norv
10	2474.2	65.2	5586	10 BC068150	BC068150 Mus muscu
11	2371	62.5	5115	4 SSC011767	AJ011767 Sus scrof
12	1939.4	51.1	2546	9 HSNUCREC	X89894 H.sapiens m
13	1485.6	39.2	1884	6 AX587466	AX587466 Sequence
14	1485.6	39.2	1884	10 AF191211	AF191211 Mus muscu
15	1435	37.8	146889	9 AL358937	AL358937 Human DNA
16	1216	32.1	1619	6 CQ723907	CQ723907 Sequence
17	1022.8	27.0	1606	4 SSC011768	AJ011768 Sus scrof
18	959.4	25.3	174140	9 AL359710	AL359710 Human DNA
19	944.8	24.9	4275	10 RNRNALRBF	X86003 R.norvegicu

20	937.4	24.7	1290	10 AF191212	AF191212 Mus muscu
21	926.6	24.4	176222	2 AC073459	AC073459 Homo sapi
22	761.4	20.1	204841	10 AL683893	AL683893 Mouse DNA
23	746.2	19.7	164637	2 AC142180	AC142180 Rattus no
24	746.2	19.7	285648	2 AC129766	AC129766 Rattus no
25	601.2	15.8	790	11 BV208714	BV208714 NR4A3_216
26	497	13.1	2790	9 BC009288	BC009288 Homo sapi
27	497	13.1	3426	6 CQ727361	CQ727361 Sequence
28	497	13.1	3427	6 A37795	A37795 Sequence 1
29	497	13.1	3427	6 AR380855	AR380855 Sequence
30	497	13.1	3427	6 AX253492	AX253492 Sequence
31	497	13.1	3427	9 HSNOT	X75918 H.sapiens m
32	496	13.1	2154	10 RNU72345	U72345 Rattus norv
33	495.4	13.1	1797	6 AX067954	AX067954 Sequence
34	492.8	13.0	2002	6 AX827766	AX827766 Sequence
35	492.8	13.0	2002	10 RU001146	U001146 Rattus norv
36	492.4	13.0	2247	10 SS3744	SS3744 NURR1-Nur-r
37	486.4	12.8	1797	6 ARI77670	ARI77670 Sequence
38	486.2	12.8	906	10 BC038520	BC038520 Mus muscu
39	483.2	12.7	2559	10 RATNR1X	L08595 Rat nuclear
40	474.2	12.5	2129	9 BC066890	BC066890 Homo sapi
41	469.4	12.4	2469	9 S77154	S77154 TINUR= NGFI
42	405.2	10.7	8805	6 AX345415	AX345415 Sequence
43	383	10.1	2169	5 BC076176	BC076176 Danio rer
44	382.2	10.1	2618	5 BC056025	BC056025 Xenopus 1
45	376.4	9.9	2738	5 BC072138	BC072138 Xenopus 1

#### ALIGNMENTS

#### RESULT 1

D78579  
LOCUS  
DEFINITION Homo sapiens NOR-1 mRNA for neuron derived orphan receptor,  
complete cds.  
ACCESSION D78579  
VERSION D78579.1 GI:1651190  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3802)  
AUTHORS Ohkura,N., Ito,M., Tsukada,T., Sasaki,K., Yamaguchi,K. and Miki,K.  
TITLE Structure, mapping and expression of a human NOR-1 gene, the third  
member of the Nur77/NGFI-B family  
JOURNAL Biochim. Biophys. Acta 1308 (3), 205-214 (1996)  
MEDLINE 96404972  
PubMed 8809112  
REFERENCE 2 (bases 1 to 3802)  
AUTHORS Ohkura,N.  
TITLE Direct Submission  
JOURNAL Submitted (30-NOV-1995) Naganari Ohkura, National Cancer Center  
Research Institute, Growth Factor Division; 5-1-1 Teukiji, Chuo-Ku,  
Tokyo 104-0045, Japan (E-mail:nohkura@ncc.ncc.go.jp,  
Tel:81-3-3542-2511(ex.4302), Fax:81-3-3542-8170)

#### FEATURES

##### source

1..3802  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="9q22"  
/tissue\_type="brain"  
/dev\_stage="fetus"

##### gene

1..3802  
/gene="NOR-1"

##### 5' UTR

1..731  
/gene="NOR-1"

##### CDS

732..2612  
/gene="NOR-1"  
/codon\_start=1



QY 1859 AACCAAGGCCCATTAACACAGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCATCT 1918  
DB 1861 AACCAAGAGCCCATTAACACAGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCATCT 1920  
QY 1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATATT 1978  
DB 1921 GCATGATGAATGCTCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATATT 1980  
QY 1979 CAGATACTGTGCCACTGACAGCGCTGCTGACGCAAGATGTCAGATGTCGCAACAAT 2038  
DB 1981 CAGATACTGTGCCACTGACAGCGCTGCTGACGCAAGATGTCAGATGTCGCAACAAT 2040  
QY 2039 TCTACAACTCTGACAGCTCCATTGATGATATCCAGAGCTGGCAGAAAGATTCGG 2098  
DB 2041 TCTACAACTCTGACAGCTCCATTGATGATATCCAGAGCTGGCAGAAAGATTCGG 2100  
QY 2099 GATTACTGATCTCCCAAGAGATCAGACATTAATTGATCAGCTTTTGGAGC 2158  
DB 2101 GATTACTGATCTCCCAAGAGATCAGACATTAATTGATCAGCTTTTGGAGC 2160  
QY 2159 TGTTCCTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTTGTGTTCTGCA 2218  
DB 2161 TGTTCCTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTTGTGTTCTGCA 2220  
QY 2219 ATGGACTTGTCTGCACTGACTTCACTGCTTCTGCTGATTTGGGGAGTGCTCGACTCTA 2278  
DB 2221 ATGGACTTGTCTGCACTGACTTCACTGCTTCTGCTGATTTGGGGAGTGCTCGACTCTA 2280  
QY 2279 TTAAGACTTTTCTTAAATTTGACAGCTGAACTTGTATATCCAGCTTACGCTGCTGCC 2338  
DB 2281 TTAAGACTTTTCTTAAATTTGACAGCTGAACTTGTATATCCAGCTTACGCTGCTGCC 2340  
QY 2339 TGTGACACTGACATGATCACAGAAAGACATGGTTAAAGAACCAAGAGAGTCAAG 2398  
DB 2341 TGTGACACTGACATGATCACAGAAAGACATGGTTAAAGAACCAAGAGAGTCAAG 2400  
QY 2399 AGCTATGCAACAGATCACAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2458  
DB 2401 AGCTATGCAACAGATCACAGCAGTTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2460  
QY 2459 AGCCACCGAGTCCAGGTCTCGGTCCCTGTAGAACTGAGAGAGATCTGCACCTCTG 2518  
DB 2461 AACCCAAAGCTCCAGGTCTCGGTCCCTGTAGAACTGAGAGAGATCTGCACCTCTG 2520  
QY 2519 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCATTTG 2578  
DB 2521 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCATTTG 2580  
QY 2579 ACAGCTCTTCTGGACACCTTACCTTCTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2638  
DB 2581 ACAAGCTCTTCTGGACACCTTACCTTCTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2640  
QY 2639 CCTCTCTAGCAGCTCTGTACGACAGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2698  
DB 2641 CCTCTCTAGCAGCTCTGTACGACAGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2700  
QY 2699 TGTCTCTCTTAAGAGGAAAAGCAGCTCTCTGTAGAAAAGCAAGACTTCTTTTTTTCTG 2758  
DB 2701 TGTCTCTCTTAAGAGGAAAAGCAGCTCTCTGTAGAAAAGCAAGACTTCTTTTTTTCTG 2760  
QY 2759 GCTCTTTCTTAAACCTAAAGCCAGAAAATTTGACAGATATTTGTTGGGTTGTGTT 2818  
DB 2761 GCTCTTTCTTAAACCTAAAGCCAGAAAATTTGACAGATATTTGTTGGGTTGTGTT 2820  
QY 2819 TTATATTAGGCATTCGGGGATGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2878  
DB 2821 TTATATTAGGCATTCGGGGATGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2880  
QY 2879 GAAATTGCTAAACAAAGCACTTTTGGACAATGCTATCCAGCAGAGAAAAGAGATAATA 2938  
DB 2881 GAAATTGCTAAACAAAGCACTTTTGGACAATGCTATCCAGCAGAGAAAAGAGATAATA 2940

QY 2939 TAACCTGTTTTAAACCTTTCTGGGAATCCAAATTAAGTGTCTTTGTATTTAAACAA 2998  
DB 2941 TAACCTGTTTTAAACCTTTCTGGGAATCCAAATTAAGTGTCTTTGTATTTAAACAA 3000  
QY 2999 GAACACCCAAAGGTTGTTCCGAGGTTAGGATGTCTTAAAGATTGGTCCCTTGAAGAT 3058  
DB 3001 GAACACCCAAAGGTTGTTCCGAGGTTAGGATGTCTTAAAGATTGGTCCCTTGAAGAT 3060  
QY 3059 ATGCTTCTGTATCAAAGGTACGTATGTGTGCAAAACAAGGCAAGAACTTCTTTAAATT 3118  
DB 3061 ATGCTTCTGTATCAAAGGTACGTATGTGTGCAAAACAAGGCAAGAACTTCTTTAAATT 3120  
QY 3119 TCTCTTCTCTTATTTTAAACAAAGTGAAGATGAGGATTTACCTACAAATCAGACAT 3178  
DB 3121 TCTCTTCTCTTATTTTAAACAAAGTGAAGATGAGGATTTACCTACAAATCAGACAT 3180  
QY 3179 GGCACAAACAATATGCTGTTTCCATAAACAAGTGAATTTTAAAGTGTCTCT 3238  
DB 3181 GGCACAAACAATATGCTGTTTCCATAAACAAGTGAATTTTAAAGTGTCTCT 3240  
QY 3239 TACTAAGTCTTGTATTAACTCTCTCTTATTTCTATATGGAATAAAGAGGAGGAGTCA 3298  
DB 3241 TACTAAGTCTTGTATTAACTCTCTCTTATTTCTATATGGAATAAAGAGGAGGAGTCA 3300  
QY 3299 TGTAGCAATGACACGTTAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3358  
DB 3301 TGTAGCAATGACACGTTAATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3360  
QY 3359 CCTTCTGAGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTG 3418  
DB 3361 CCTTCTGAGTATGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTG 3420  
QY 3419 CCTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTAATAACTGGATTACATGTAAGT 3478  
DB 3421 CCTGACTGTCCAGCTATCCTGAAAGTGGATCAGATTAATAACTGGATTACATGTAAGT 3480  
QY 3479 TTTTGGTGTGTCTTATCAACCCACAGAGTTCCTTAACTTGTCTTCAAGTATAGTAAC 3538  
DB 3481 TTTTGGTGTGTCTTATCAACCCACAGAGTTCCTTAACTTGTCTTCAAGTATAGTAAC 3540  
QY 3539 TGACTGTATATTTTATTTCATTGAGAGCGCATTAAGTTCAGTATTTGATCCCTAGATAAG 3598  
DB 3541 TGACTGTATATTTTATTTCATTGAGAGCGCATTAAGTTCAGTATTTGATCCCTAGATAAG 3600  
QY 3599 AACATGCAAAATCAGCAGGAACCTGGTCAATACAGGTAAGCACCAGGCAATAAGGATTTT 3658  
DB 3601 AACATGCAAAATCAGCAGGAACCTGGTCAATACAGGTAAGCACCAGGCAATAAGGATTTT 3660  
QY 3659 TATAGATATAATTTAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3715  
DB 3661 TATAGATATAATTTAATTTTGTTAATTTGGTTAAGGAGACCAATTTTGGAGAGCAAGCA 3720  
QY 3716 AA---TCTTTTAAAAAATAGTATGAATCTAGAAAAGATTTAAAAAATAGTAT 3772  
DB 3721 AATCTTCTTTTAAAAAATAGTATGAATCTAGAAAAGATTTAAAAAATAGTAT 3780  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
DB 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 2  
E12854  
LOCUS  
DEFINITION  
cDNA encoding neuron derived orphan receptor, NOR-1.  
E12854  
ACCESSION  
E12854.1 GI:3251686  
VERSION  
JP 1997084585-A/1.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 3762)





QY 1659 GGGCTGCAAGGGCTTTTCAAGAGACAGTGCAGAAAAATGCAAAATATATGTTGCTGGC 1718  
DB 1621 GGGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAAAAATGCAAAATATATGTTGCTGGC 1680  
QY 1719 AAATAAAAACTGCCAGTAGACAAGAGAGAGTGCAGAACCGATGTCAGTACTGTCGATTTCA 1778  
DB 1681 AAATAAAAACTGCCAGTAGACAAGAGAGAGTGCAGAACCGATGTCAGTACTGTCGATTTCA 1740  
QY 1779 GAAGTGTCTCAGTGTGGAAATGTTAAAGAGTGTGTCGATACAGATAGTCTGAAGGGAG 1838  
DB 1741 GAAGTGTCTCAGTGTGGAAATGTTAAAGAGTGTGTCGATACAGATAGTCTGAAGGGAG 1800  
QY 1839 GAGAGTGTCTGCTCTCCAAACCAAGAGCCATTACACAGGACCTTCTCAGCCCTC 1898  
DB 1801 GAGAGTGTCTGCTCTCCAAACCAAGAGCCATTACACAGGAACTTCTCAGCCCTC 1860  
QY 1899 TCCACCTTCTCCTCCAAATCTGCATGATGAATGCCCTTGTCCGAGCTTTTAACAGACTCAAC 1958  
DB 1861 TCCACCTTCTCCTCCAAATCTGCATGATGAATGCTCTTGTCCGAGCTTTTAACAGACTCAAC 1920  
QY 1959 ACCAGAGATCTTGATTTTCCAGATACCTGTCCTCACTGACACAGGCTGCTGCAGGCACAGA 2018  
DB 1921 ACCAGAGATCTTGATTTTCCAGATACCTGTCCTCACTGACACAGGCTGCTGCAGGCACAGA 1980  
QY 2019 TGCTGAGCATGTGCACAAATTTCTACAACTCTGACAGCCTCCATTGATGTATCCAGAAG 2078  
DB 1981 TGCTGAGCATGTGCACAAATTTCTACAACTCTGACAGCCTCCATTGATGTATCCAGAAG 2040  
QY 2079 CTGGGCAGAAAGATTTCCGGGATTTACTGATCTCCCAAGAGATGACAGACTTACTTAT 2138  
DB 2041 CTGGGCAGAAAGATTTCCGGGATTTACTGATCTCCCAAGAGATGACAGACTTACTTAT 2100  
QY 2139 TGAATCAGCCTTTTGGAGCTGTTGCTCTCAGACTTTCCATCAGGTCAAAACACTGCTGA 2198  
DB 2101 TGAATCAGCCTTTTGGAGCTGTTGCTCTCAGACTTTCCATCAGGTCAAAACACTGCTGA 2160  
QY 2199 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCACTGACATTCAGTGCCTTCGTGGATT 2258  
DB 2161 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCACTGACATTCAGTGCCTTCGTGGATT 2220  
QY 2259 TGGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAATTTGACAGGCTGAACCTTGA 2318  
DB 2221 TGGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAATTTGACAGGCTGAACCTTGA 2280  
QY 2319 TATCCAAGCCTTAGCCTGCTGTGACACTGAGCATGATCACAGAAAGACATGGGTAAA 2378  
DB 2281 TATCCAAGCCTTAGCCTGCTGTGACACTGAGCATGATCACAGAAAGACATGGGTAAA 2340  
QY 2379 AGAACCAAGAGAGTGCAGAGCTATGCAACAGATCACAGGAGTTTAAAGACCACCA 2438  
DB 2341 AGAACCAAGAGAGTGCAGAGCTATGCAACAGATCACAGGAGTTTAAAGACCACCA 2400  
QY 2439 GAGTAAGGACAGGCTCTGGAGCCACCGAGTCCCAAGGCTCTGGGTGCCCTGTTGAACT 2498  
DB 2401 GAGTAAGGACAGGCTCTGGAAACCAAGAGTCCCAAGGCTCTGGGTGCCCTGTTGAACT 2460  
QY 2499 GAGGAAGATCTGCACCCCTGGGCCCTCCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGGT 2558  
DB 2461 GAGGAAGATCTGCACCCCTGGGCCCTCCAGGCGATCTTCTACCTGAAGCTGGAAGACTTGGT 2520  
QY 2559 GTCTCCACCTTCATCTGACAGCTCTTCTGGAACAGCTTCTTACCTGAGGAGC 2618  
DB 2521 GTCTCCACCTTCATCTGACAGCTCTTCTGGAACAGCTTCTTACCTGAGGAGC 2580  
QY 2619 AGTGGAGCAGTGAAGCTCTCTCTAGCAGCTCTGCTGACGAGCAAAAGGGATAGS 2678  
DB 2581 AGTGGAGCAGTGAAGCTCTCTCTAGCAGCTCTGCTGACGAGCAAAAGGGATAGS 2640  
QY 2679 TTTGGAACCTATCTATTTCTCTCTCTTAAAGAGAAAAGCAGCTCTCTGTAGAAGCA 2738  
DB 2641 TTTGGAACCTATCTATTTCTCTCTCTTAAAGAGAAAAGCAGCTCTCTGTAGAAGCA 2700

QY 2739 AAGACTTTCTTTTCTGCTCTTTTCTTACACCTAAGCCAGAAAACCTTGCGAG 2798  
DB 2701 AAGACTTTCTTTTCTGCTCTTTTCTTACACCTAAGCCAGAAAACCTTGCGAG 2760  
QY 2799 TATTGTGTTGGGTTGTGTTTATATTAGGCATTTGGGGATGGGGTGGAGGGGTAT 2858  
DB 2761 TATTGTGTTGGGTTGTGTTTATATTAGGCATTTGGGGATGGGGTGGAGGGGTAT 2820  
QY 2859 AGTTCATGAGGGTTTCTAAGAAAATGCTAACAAAGCACCTTTTGGCAATGCTATCCCAG 2918  
DB 2821 AGTTCATGAGGGTTTCTAAGAAAATGCTAACAAAGCACCTTTTGGCAATGCTATCCCAG 2880  
QY 2919 CAGGAAAAAAGGATAATAAATCTTTTAAAACTCTTTCTGGGGAATCCAAATATAGT 2978  
DB 2881 CAGGAAAAAAGGATAATAAATCTTTTAAAACTCTTTCTGGGGAATCCAAATATAGT 2940  
QY 2979 TGTCTTGATTTTAAAAAAGAACAGCAGCAGGGTGTTCGCCAGGGTAGGATGTCTTA 3038  
DB 2941 TGTCTTGATTTTAAAAAAGAACAGCAGCAGGGTGTTCGCCAGGGTAGGATGTCTTA 3000  
QY 3039 AAGATTGTCCTTGAAAATATGCTTCTGTATCAAGGTACGTATGTGTGCAAAACAG 3098  
DB 3001 AAGATTGTCCTTGAAAATATGCTTCTGTATCAAGGTACGTATGTGTGCAAAACAG 3060  
QY 3099 GCAGAAACTTCTTTTAAATTTCTCTTCTTTTAACTTAACTGTAAGAGTGGAGG 3158  
DB 3061 GCAGAAACTTCTTTTAAATTTCTCTTCTTTTAACTTAACTGTAAGAGTGGAGG 3120  
QY 3159 ATTACTACAAATCAGACATGCAAAACAAATATGCTGTTCCTTCCATAAACAAAGTGC 3218  
DB 3121 ATTACTACAAATCAGACATGCAAAACAAATATGCTGTTCCTTCCATAAACAAAGTGC 3180  
QY 3219 AATTTTTTAAAGTGTCTTACTAAGTCTGTGTTTAACTCTCTCTTTTATTCTATATGG 3278  
DB 3181 AATTTTTTAAAGTGTCTTACTAAGTCTGTGTTTAACTCTCTCTTTTATTCTATATGG 3240  
QY 3279 AATAAAAGAGGAGGAGTCATGTTAGCAATGACAGCTTAAATCTCTAGCAGAGCTGT 3338  
DB 3241 AATAAAAGAGGAGGAGTCATGTTAGCAATGACAGCTTAAATCTCTAGCAGAGCTGT 3300  
QY 3339 GTTACCTTCCCTGTGCTGATCCCTTCTGAGGTATGGCCCATCCAAAGACTTTTAGGCCATTC 3398  
DB 3301 GTTACCTTCCCTGTGCTGATCCCTTCTGAGGTATGGCCCATCCAAAGACTTTTAGGCCATTC 3360  
QY 3399 TTGATGAAACAGATCCCTGCCCTGACTGTCCAGCTATCCTGAAAGTGAATCAGATTATA 3458  
DB 3361 TTGATGAAACAGATCCCTGCCCTGACTGTCCAGCTATCCTGAAAGTGAATCAGATTATA 3420  
QY 3459 AACTGGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCACCCAGAGTTCCCTAAA 3518  
DB 3421 AACTGGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCACCCAGAGTTCCCTAAA 3480  
QY 3519 CTTGCTTCAGTTATAGTAACTGACTGTTATATTTCATTAGAAAGCGCCATAAGTCAGTTGA 3578  
DB 3481 CTTGCTTCAGTTATAGTAACTGACTGTTATATTTCATTAGAAAGCGCCATAAGTCAGTTGA 3540  
QY 3579 GTATTGATCCCTAGATAAGAACTGCAAAATCAGCAGAACTGGGTCAATCAGGGTAAGCA 3638  
DB 3541 GTATTGATCCCTAGATAAGAACTGCAAAATCAGCAGAACTGGGTCAATCAGGGTAAGCA 3600  
QY 3639 CCAGGACAAATAGGATTTTATAGATATAATTTAAATTTTGTTA--TTGGTTAAGGAGA 3696  
DB 3601 CCAGGACAAATAGGATTTTATAGATATAATTTAAATTTTGGTAATGGGTTAAGGAGA 3660  
QY 3697 -CAATTTTGGAGCAAGCAAA---TCTTTTTTAAAAAATAGTATGATGTAATCTAGTA 3752  
DB 3661 CCAATTTTGGAGGCAAGCAAACTCTTTTAAAAAATAGTATGTAATGTAATCTAGTA 3720  
QY 3753 AAAGATTTTAAAAAATAGTATGATGAGTACTAGGAAGGAT 3794  
DB 3721 AAAGATTTTAAAGAAATAGTATGATGAGTACTAGGAAGGAT 3762





3'UTR	1974..4977	75.8%; Score 2875; DB 9; Length 4977;	
polyA_signal	/gene="MINOR"	Best Local Similarity 98.1%; Pred. No. 0;	
	4953..4958	Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;	
polyA_site	/gene="MINOR"		
	4977		
	/gene="MINOR"		
	/note="8 A residues"		
ORIGIN			
Query Match			
Best Local Similarity			
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;			
QY	733 CCTCGTCCCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790		
DB	94 CCTGTCGTCGAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCAGA 153		
QY	791 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGTACACCAAGCTGACATGG 850		
DB	154 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGTACACCAAGCTGACATGG 213		
QY	851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACACGTCCTGCTGCCAGCATCAGTACT 910		
DB	214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACACGTCCTGCTGCCAGCATCAGTACT 273		
QY	911 TCGTGGAGGGTACTCGAGCAACTAGAACTCAAGCCTTCTGCGTGACCAAAATGACAGC 970		
DB	274 TTGTGGAGGGTACTCGAGCAACTAGAACTCAAGCCTTCTGCGTGACCAAAATGACAGC 333		
QY	971 GGCCTTGATCAAAGTGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1030		
DB	334 GGCCTTGATCAAAGTGAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 393		
QY	1031 ACCAC 1090		
DB	394 ACCAC 453		
QY	1091 CCAGCCCGAGGAGCAGGTGCTGCCAGCACCTCCATGCTTCAAGCAGTCCCAACCGT 1150		
DB	454 CCAGCCCGAGGAGCAGGTGCTGCCAGCACCTCCATGCTTCAAGCAGTCCCAACCGT 513		
QY	1151 CCACCCCAAC 1210		
DB	514 CCACCCCAAC 573		
QY	1211 CCTCGCGCCGCGTGCATCGACACCGCGCGCTGTCGACCGCGCGCGTGAAGGCGGTCC 1270		
DB	574 CCTCGCGCCGCGTGCATCGACACCGCGCGCTGTCGACCGCGCGCGTGAAGGCGGTCC 633		
QY	1271 CCACGGTGGCGCGCGCGCTTCCGCGCTTCCACTTCAAGCCCTCGCGCGCGCATCCCC 1330		
DB	634 CCACGGTGGCGCGCGCGCTTCCGCGCTTCCACTTCAAGCCCTCGCGCGCGCATCCCC 693		
QY	1331 CGCGCCCAACCG 1390		
DB	694 CGCGCCCAACCG 753		
QY	1391 TCAGCCTGCGCTGGAGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1450		
DB	754 TCAGCCTGCGCTGGAGCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 813		
QY	1451 ACCCGTACGGGTGCGCTGGCCAAAGAGGGCGGGCCCGCTGCGCTTCCGCGCTTCCGCG 1510		
DB	814 ACCCGTACGGGTGCGCTGGCCAAAGAGGGCGGGCCCGCTGCGCTTCCGCGCTTCCGCG 873		
QY	1511 TCAGCCTTCCCTACCGGCTCAGGCTCGGCGAGAGTCCAGCCTGCGCGCTGCGCGCG 1570		
DB	874 TCAGCCTTCCCTACCGGCTCAGGCTCGGCGAGAGTCCAGCCTGCGCGCTGCGCGCG 933		
QY	1571 CCAGCAGAGCTCGCTGCTGCGGAGGGGACGCTGTGCGGCTGTGCGGGGACACCGCGCT 1630		
DB	934 CCAGCAGAGCTCGCTGCTGCGGAGGGGACGCTGTGCGGCTGTGCGGGGACACCGCGCT 993		
QY	1631 GCCAGCACTACGGCTGCGAACCCTCGAGGGGTGCAAGGGCTTTTTCAGAGAACAGTGC 1690		
DB	994 GCCAGCACTACGGCTGCGAACCCTCGAGGGGTGCAAGGGCTTTTTCAGAGAACAGTGC 1053		
QY	1691 AGAAAAATGCAAAATATGTTTGCCTGGCAATAAAACTGCCAGTACAGACAGACGTC 1750		
DB	1054 AGAAAAATGCAAAATATGTTTGCCTGGCAATAAAACTGCCAGTACAGACAGACGTC 1113		
QY	1751 GAAACCGATGTGCTGATTTCAAGAGTGTCTCAGTGTGGAATGGTAAAGAGAG 1810		
DB	1114 GAAACCGATGTGCTGATTTCAAGAGTGTCTCAGTGTGGAATGGTAAAGAGAG 1173		
QY	1811 TTGTCGCTACAGATAGTCTGAAAGGAGGAGAGTGTGCTGCTTCCCAAAACCAAGAGCC 1870		
DB	1174 TTGTCGCTACAGATAGTCTGAAAGGAGGAGAGTGTGCTGCTTCCCAAAACCAAGAGCC 1233		
QY	1871 CATTAACAAGGAACCTTCTCAGCCCTCTCCACTTCTCCTCCAATCTGCATGATGAATG 1930		
DB	1234 CATTAACAAGGAACCTTCTCAGCCCTCTCCACTTCTCCTCCAATCTGCATGATGAATG 1293		
QY	1931 CCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATTCAGATACTGTC 1990		
DB	1294 CTCTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTTATTCAGATACTGTC 1353		
QY	1991 CCACTGACAGGCTGCTGACGACAGATGCTGAGCATGTGCAACAAATCTACACCTCC 2050		
DB	1354 CCACTGACAGGCTGCTGACGACAGATGCTGAGCATGTGCAACAAATCTACACCTCC 1413		
QY	2051 TGACAGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCGGGATTTACTGATC 2110		
DB	1414 TGACAGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCGGGATTTACTGATC 1473		
QY	2111 TCCCAAAAGAGATCAGACATTTACTTATTAATCAGCCCTTTTGGAGCTGTTTGTCTCTCA 2170		
DB	1474 TCCCAAAAGAGATCAGACATTTACTTATTAATCAGCCCTTTTGGAGCTGTTTGTCTCTCA 1533		
QY	2171 GACTTTCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTTGCATGAGCTTGTCC 2230		
DB	1534 GACTTTCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTTGCATGAGCTTGTCC 1593		
QY	2231 TGCACTGACTTCAGTGCCTTCTGATTTGGGAGTGGCTGACTCTATAAGACTTTT 2290		
DB	1594 TGCACTGACTTCAGTGCCTTCTGATTTGGGAGTGGCTGACTCTATAAGACTTTT 1653		
QY	2291 CCTTAAATTTGCAGAGCCTGAACCTTGATATCCAAAGCCTTAGCCTGCCTGTGACGACTGA 2350		
DB	1654 CCTTAAATTTGCAGAGCCTGAACCTTGATATCCAAAGCCTTAGCCTGCCTGTGACGACTGA 1713		
QY	2351 GCATGATCACAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACA 2410		
DB	1714 GCATGATCACAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACA 1773		
QY	2411 AGATCACAGCAGTTTAAAGAACCAAGAGTGAAGAGCAGCTCTGGAGCCACCGAGT 2470		
DB	1774 AGATCACAGCAGTTTAAAGAACCAAGAGTGAAGAGCAGCTCTGGAGCCACCGAGT 1833		
QY	2471 CCAAGCTCCTGGGTGCCCTGTGAGAACTGAGGAAGATCTGCAACCTGGGCTTCCAGCGCA 2530		
DB	1834 CCAAGCTCCTGGGTGCCCTGTGAGAACTGAGGAAGATCTGCAACCTGGGCTTCCAGCGCA 1893		
QY	2531 TCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACTTCCATCATTTGACAGACTTTC 2590		
DB	1894 TCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACTTCCATCATTTGACAGACTTTC 1953		
QY	2591 TGGACACCTACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTCTCTCTCTCTCT 2650		
DB	1954 TGGACACCTACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTCTCTCTCTCTCT 2013		
QY	2651 CCTGCTTCTACGCAAGAAAGGATAGGTTTGGAAACCTTATCATTTCTCTCTCTCTCTCT 2710		
DB	2014 CCTGCTTCTACGCAAGAAAGGATAGGTTTGGAAACCTTATCATTTCTCTCTCTCTCTCT 2073		
QY	2711 AGAGAAAGACAGCTCTCTGTAGAAAGAAAGACTTTCTTTTCTTTTCTTTTCTTTTCTTT 2770		

```
Db 2074 AGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTCTGCTCTTTTCCTT 2133
QY 2771 ACAACCTAAAGCCAGAAAACTTGCAGAGTATTGTGTGGGTGTGTTTTATATTAGGC 2830
Db 2134 ACAACCTAAAGCCAGAAAACTTGCAGAGTATTGTGTGGGTGTGTTTTATATTAGGC 2193
QY 2831 ATTGGGGATGGGTGGGAGGGGTATAGTTTCATGAGGGTTTTCTAAGAAATTCCTAAC 2890
Db 2194 ATTGGGGATGGGTGGGAGGGGTATAGTTTCATGAGGGTTTTCTAAGAAATTCCTAAC 2253
QY 2891 AAAGCACATTTCGACAATCTATCCAGCAGGAAAAAAGGATATATAACTGTGTTTAA 2950
Db 2254 AAAGCACATTTCGACAACTCTATCCAGCAGGAAAAAAGGATATATAACTGTGTTAA 2313
QY 2951 AACTCTTCTGGGGAATCAATATATAGTTGCTTTGATTTAAACAAAGAACAGCAAGG 3010
Db 2314 AACTCTTCTGGGGAATCAATATATAGTTGCTTTGATTTAAACAAAGAACAGCAAGG 2373
QY 3011 GTTGTTCGACAGGTAGGATGTCTTAAAGATTGCTCCCTTGAAAAATATGCTTCTGTA 3070
Db 2374 GTTGTTCGACAGGTAGGATGTCTTAAAGATTGCTCCCTTGAAAAATATGCTTCTGTA 2433
QY 3071 TCAAGGTACGTATGTGGTGCACAAAGCAGAACTTCCTTTTAAATTCCTTCTTCCTT 3130
Db 2434 TCAAGGTACGTATGTGGTGCACAAAGCAGAACTTCCTTTTAAATTCCTTCTTCCTT 2493
QY 3131 TATTTTAAACAAATGTGAAAGATGAGGATTACCTACAATCAGACATGCGCAAAACAATA 3190
Db 2494 TATTTTAAACAAATGTGAAAGATGAGGATTACCTACAATCAGACATGCGCAAAACAAT- 2551
QY 3191 ATGGCTGTTTGTTCATATAACAAAGTGCAATTTTTTAAAGTGTGCTTAAAGTCTTG 3250
Db 2552 ATGGCTGTTTGTTC---ATAACAGCTGCAATTTTTTAAAGTGTGCTTAAAGTCTTG 2608
QY 3251 TTTTATAACTCTCCTTTATTCTATATGGAATAAAGAGGAGCAGTCATGTTAGCAATG 3310
Db 2609 T---TATACTCTCTTATCTATATGCGGAAATAAAGAGGAGCAGTCATGTTAGCAATG 2665
QY 3311 ACACGTTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCTCTGCATCCCTTCTGAGSTA 3370
Db 2666 ACACGTTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCTCTGCATCCCTTCTGAGSTA 2725
QY 3371 TGGCCCATCCAGACTTTTATAGCCATCTTTGATGGAACAGATCCCTGCTGACTGTCC 3430
Db 2726 TGGCCCATCCAGACTTTTATAGCCATCTTTGATGGAACAGATCCCTGCTGACTGTCC 2785
QY 3431 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGTGTTGTTGT 3490
Db 2786 AGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGTGTTGTTGT 2845
QY 3491 TCTATCAACCCACACAGAGTTCCTTAACTTGCTTCAGTTATAGTAAGTGTGTTATAT 3550
Db 2846 TCTATCAACCCACACAGAGTTCCTTAACTTGCTTCAGTTATAGTAAGTGTGTTATAT- 2904
QY 3551 TCATTAGAGGCCCATTAAGTCAAGTTGATTTGATCCCTAGATAAGAACATGCAATC 3610
Db 2905 TCATTAGAGGCCCATTAAGTCAAGTTGATTTGATCCCTAGATAAGAACATGCAATC--- 2959
QY 3611 AGCAGGAAGTGTATACAGGTTAGCACACAGGCAATTAAGATTTTATAGATATAT 3670
Db 2960 TCAGAGGAGTGTGTATACAGGTTAGCACACAGGCAATTAAGATTTTATAGATATAT 3019
QY 3671 TTAATTTTGTATTGTTTAAAGAGACAATTTTGGAGACCAAGCAAA---TCTTTTAA 3727
Db 3020 TTAATTTTGTATTGTTTAAAGAGACAATTTTGGAGACCAAGCAAACTCTCTTTTAA 3079
QY 3728 AAATAGTATCAATGTGAATCTAGAAAGATTTTAAATAATAGTATGAGTGTGAGTACTAG 3787
Db 3080 AAATAGTATCAATGTGAATCTAGAAAGATTTTAAATAATAGTATGAGTGTGAGTACTAG 3139
QY 3788 GAAGGAT 3794
|||||
```

```
Db 3140 GAAGGAT 3146
RESULT 5
AR380383
LOCUS AR380383 2714 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 928 from patent US 6607879.
ACCESSION AR380383
VERSION AR380383.1 GI:40088017
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2714)
AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
TITLE Compositions for the detection of blood cell and immunological
response gene expression
JOURNAL Patent: US 6607879-A 928 19-AUG-2003;
FEATURES
LOCATION/Qualifiers
source 1..2714
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 70.9%; Score 2689.6; DB 6; Length 2714;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 132 TCACCTCGCACACACAGACACAGCGCGACACAGGCTCCGCACACACACTTCGCTCTCCC 191
Db 19 TCACACAAACACACAGACACACAGCGGCACACAGGCTCCGCACACACACTTCGCTCTCCC 78
QY 192 GCGCGCTCACACCCCTCTTGGCCCTGAGCCCTTTCGCGGTGACGCGCGCGCGAGCTGGA 251
Db 79 GCGCGCTCACACCCCTCTTGGCCCTGAGCCCTTTCGCGGTGACGCGCGCGCGAGCTGGA 138
QY 252 CGCCCTCCCGGGCTCACCTTTCGAACGCTGACGGTCCGCGAGTGGCCGTGGAGGTGGGA 311
Db 139 CGCCCTCCCGGGCTCACCTTTCGAACGCTGACGGTCCGCGAGTGGCCGTGGAGGTGGGA 198
QY 312 ACAGCGGGGACATCTCCCGCTGGTACAGCCCAAGCCAGGACGCGCGGAAACCTCTC 371
Db 199 ACAGCGGGGACATCTCCCGCTGGTACAGCCCAAGCCAGGACGCGCGGAAACCTCTC 258
QY 372 GGTGTGCTCTCCATGAGTCGGGATCGCAGATCCCCCAGCGCTCACGCGCTCCG 431
Db 259 GGTGTGCTCTCCATGAGTCGGGATCGCAGATCCCCCAGCGCTCACGCGCTCCG 318
QY 432 GGAGCGCTGGGCTTGTACACCGAGCCCTTCGGGACAGCAGCTGTGACTCCCCCCCAG 491
Db 319 GGAGCGCTGGGCTTGTACACCGAGCCCTTCGGGACAGCAGCTGTGACTCCCCCCCAG 378
QY 492 TGCAGATTCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAACCGCCACAGCACTC 551
Db 379 TGCAGATTCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAACCGCCACAGCACTC 438
QY 552 AAAGCCCACTGCGGAAGAGGGCAGCCCGCAAGCCCGGCCCTGAGCCCTGGACCTTAGC 611
Db 439 AAAGCCCACTGCGGAAGAGGGCAGCCCGCAAGCCCGGCCCTGAGCCCTGGACCTTAGC 498
QY 612 GGTGCGGCGACACTGCGCGGCTTTCGCTCGCGGACGTCGCTTCCTTACACTCTC 671
Db 499 GGTGCGGCGACACTGCGCGGCTTTCGCTCGCGGACGTCGCTTCCTTACACTCTC 558
QY 672 AGCCTCGCTGAGAGAGACCCCGACCCCACTTCAGCGCGCAAGATACCTCCACATAT 731
Db 559 AGCCTCGCTGAGAGAGACCCCGACCCCACTTCAGCGCGCAAGATACCTCCACATAT 618
QY 732 GGCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGGTTATGCGCGCAGAC 791
Db 619 GGCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGGTTATGCGCGCAGAC 678
QY 792 ATACAGCTCGGAATACACACGAGATATATGAACCCCGACTACACCAAGCTGACCATGGA 851
```

679	Db	 ATACAGCTCGGAATACACCA	 CGGAGATCATGAAC	 CCCGAGCTACACCA	 AGGTGACCATGGA	738
852	Qy	 CCTTGGCAGCACTCAGAT	 CAGCGCTCAGACCA	 CACCACGCTCCTG	 TCCGCCAGCATCAGTACCTT	911
739	Db	 CCTTGGCAGCACTCAGAT	 CAGCGCTCAGACCA	 CCACGCTCCTG	 TCCGCCAGCATCAGTACCTT	798
912	Qy	 CGTGGAGGGCTACTCGAG	 CAACTCAGAACTCAAG	 CCTTCTGCGTGTACCA	 AATGCGAGCG	971
799	Db	 CGTGGAGGGCTACTCGAG	 CAACTCAGAACTCAAG	 CCTTCTGCGTGTACCA	 AATGCGAGCG	858
972	Qy	 GCCCTTGATCAAAGTGA	 GAGAGGGCGCGCCAG	 CTACATCAATCACA	 CACACACACCA	1031
859	Db	 GCCCTTGATCAAAGTGA	 GAGAGGGCGCGCCAG	 CTACATCACCATCA	 CACCACCA	918
1032	Qy	 CCACCACCAACCACCAT	 CACCAGCAGCAGCAT	 CAGCAGCCATCC	 ATTCCTCCAGCCTC	1091
919	Db	 CCACCACCAACCACCAT	 CACCAGCAGCAGCAT	 CAGCAGCCATCC	 ATTCCTCCAGCCTC	978
1092	Qy	 CAGCCCGGAGGACGAGGT	 GTGCTGCCACGACCT	 CCATGTCTTCAAG	 CAGTCCCAACCGCTC	1151
979	Db	 CAGCCCGGAGGACGAGGT	 GTGCTGCCACGACCT	 CCATGTCTTCAAG	 CAGTCCCAACCGCTC	1038
1152	Qy	 CACCCCAACAACCGCGCT	 TTCCTCCCGCAGCGGG	 GGCGTTATGGA	 CGAGGCACTGCC	1211
1039	Db	 CACCCCAACAACCGCGCT	 TTCCTCCCGCAGCGGG	 GGCGTTATGGA	 CGAGGCACTGCC	1098
1212	Qy	 CTGGCGCCCGCTGCAT	 CCGACCGCGCGCTGT	 GACCCCGATGAAG	 CGGGTCCC	1271
1099	Db	 CTGGCGCCCGCTGCAT	 CGACCGCGCGCGCTGT	 GGAACCGCGATGAAG	 CGGGTCCC	1158
1272	Qy	 CAGCGTGGCGCGCGCGCT	 TCCCGCTTCCACTTCA	 AGCCCTCGCGCGCAT	 CCCC	1331
1159	Db	 CAGGTGGCGCGCGCGCT	 TCCCGCTTCCACTTCA	 AGCCCTCGCGCGCAT	 CCCC	1218
1332	Qy	 CGCGCCACAGCCCGCGC	 CGGCGCACACCTCGGT	 ACGAACCGACGCGCT	 CGCGCT	1391
1219	Db	 CGCGCCACAGCCCGCGC	 CGGCGCACACCTCGGT	 ACGAACCGACGCGCT	 CGCGCT	1278
1392	Qy	 CAGCCTCGCCTGGAGCG	 CAGCCGCGGCGCAGCG	 CGCGCGCTTCAAG	 AGCCA	1451
1279	Db	 CAGCCTCGCCTGGAGCG	 CAGCGCGCGCGGAGC	 CAGCGCGCGCTTCAAG	 AGCCA	1338
1452	Qy	 CCCGTACGGGTGCGCGT	 TGCGCAAGAGGGCGG	 CGCCCTCGCGCTCT	 CGCGCT	1511
1339	Db	 CCCGTACGGGTGCGCGT	 TGCGCAAGAGGGCGG	 CGCCCTCGCGCTCT	 CGCGCT	1398
1512	Qy	 CAGCCCTCCCTACCGCG	 TCAAGCTGTGGGCGAG	 AGTCCAGCGCTCGCG	 CGCG	1571
1399	Db	 CAGCCCTCCCTACCGCG	 TCAAGCTGTGGGCGAG	 AGTCCAGCGCTCGCG	 CGCG	1458
1572	Qy	 CAGCAGAGCTCGTCTGT	 GCGAGGCACTGTGCG	 TGTGCGGGGACAA	 CGCGCTG	1631
1459	Db	 CAGCAGAGCTCGTCTGT	 GCGAGGCACTGTGCG	 TGTGCGGGGACAA	 CGCGCTG	1518
1632	Qy	 CCAGCACTACGGCGTGA	 CAACTCGAGGGCTGCA	 AGGGCTTTTCAAG	 AGAACAGTGA	1691
1519	Db	 CCAGCACTACGGCGTGA	 CAACTCGAGGGCTGCA	 AGGGCTTTTCAAG	 AGAACAGTGA	1578
1692	Qy	 GAAAAATGCAAAATATG	 TTTTCCTGGCAAAATA	 AAAACTGCCCACTAGA	 CAAGAGACGTG	1751
1579	Db	 GAAAAATGCAAAATATG	 TTTTCCTGGCAAAATA	 AAAACTGCCCACTAGA	 CAAGAGACGTG	1638
1752	Qy	 AAACCGATGTCAGTACT	 GTGCGATTCAGAGTGT	 CTCAGTGTGGATGT	 TGTAAAGAGT	1811
1639	Db	 AAACCGATGTCAGTACT	 GTGCGATTCAGAGTGT	 CTCAGTGTGGATGT	 TGTAAAGAGT	1698
1812	Qy	 TGTCGCTACAGATAGTCT	 GAAAGGAGAGAGTGT	 CTGCTTCCAAAC	 CCAAAGAGCCC	1871
1699	Db	 TGTCGCTACAGATAGTCT	 GAAAGGAGAGAGTGT	 CTGCTTCCAAAC	 CCAAAGAGCCC	1758
1872	Qy	 ATTACAAAGGAACCTTCT	 CAGCCCTCTCCACCTT	 CTCTCCAATTCGAT	 TGATGATGC	1931

Db	1759	ATTACAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAACTGCGCATGATGAATGC	1818
Qy	1932	CTTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGACTTTGATTTATTTCCAGATACTGTCC	1991
Db	1819	CTTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGACTTTGATTTATTTCCAGATACTGTCC	1878
Qy	1992	CATGACACAGGCTCTCGAGGACAGATGCTGAGCATGTGCGAACAAATTCACAACTTCCT	2051
Db	1879	CATGACACAGGCTCTCGAGGACAGATGCTGAGCATGTGCGAACAAATTCACAACTTCCT	1938
Qy	2052	GACAGCTCCATTGATGTATCCAGAACTGGGCGAAGAAATTCGCGGATTTACTGATCT	2111
Db	1939	GACAGCTCCATTGATGTATCCAGAACTGGGCGAAGAAATTCGCGGATTTACTGATCT	1998
Qy	2112	CCCCAAAGAGATCAGACATTTACTTTATTTGAATCAGCCCTTTTGGAGCTGTTTGTGCTTCAG	2171
Db	1999	CCCCAAAGAGATCAGACATTTACTTTATTTGAATCAGCCCTTTTGGAGCTGTTTGTGCTTCAG	2058
Qy	2172	ACTTTCCATCAGGTTCAACACATCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTTGCTCT	2231
Db	2059	ACTTTCCATCAGGTTCAACACATCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTTGCTCT	2118
Qy	2232	GCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTTC	2291
Db	2119	GCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTTC	2178
Qy	2292	CTTAAATTTGACAGAGCTGAACTTTGATATTCCAAGCCTTAGCCCTGCCTGTGACGACTGAG	2351
Db	2179	CTTAAATTTGACAGAGCTGAACTTTGATATTCCAAGCCTTAGCCCTGTGACGACTGAG	2238
Qy	2352	CATGATCACAGAAAGACATGGTGTAAAGAAACCAAGAGAGTGGAGAGCTATGCAACAA	2411
Db	2239	CATGATCACAGAAAGACATGGTGTAAAGAAACCAAGAGAGTGGAGAGCTATGCAACAA	2298
Qy	2412	GATCACAGAGCTTTAAAGACACACAGAGTAAGGACAGGCTCTGAGAGCCACCGAGTC	2471
Db	2299	GATCACAGAGCTTTAAAGACACACAGAGTAAGGACAGGCTCTGAGAGCCACCGAGTC	2358
Qy	2472	CAAGGCTCTGGTGCCCTGGTGTAGAACTGAGAGAGCTCGACCCCTGGGCTCCAGAGCCAT	2531
Db	2359	CAAGGCTCTGGTGCCCTGGTGTAGAACTGAGAGAGCTCGACCCCTGGGCTCCAGAGCCAT	2418
Qy	2532	CTTCTACTGAACTGGAAGACTTTGGTGTCTCCAACCTTCCATCATGCAAGCTCTTCTCT	2591
Db	2419	CTTCTACTGAACTGGAAGACTTTGGTGTCTCCAACCTTCCATCATGCAAGCTCTTCTCT	2478
Qy	2592	GGACACCTTACTTTCTAATCAGAGAGCTGGAGAGTGGAGCTGCTCTCTCTCTAGCAC	2651
Db	2479	GGACACCTTACTTTCTAATCAGAGAGCTGGAGAGTGGAGCTGCTCTCTCTCTAGCAC	2538
Qy	2652	CTGCTTGTCTACGACAGAAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTTCCCTTAA	2711
Db	2539	CTGCTTGTCTACGACAGAAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTTCCCTTAA	2598
Qy	2712	GAGGAAAGCAGCTCTCTGTAGAAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCTCTTA	2771
Db	2599	GAGGAAAGCAGCTCTCTGTAGAAAGCAAGACTTTCTTTTTTCTGGCTCTTTTCTCTTA	2658
Qy	2772	CAACCTAAAGCCAGAAACCTTCGAGAGTATTTGTTGGGGTGTGTTTTATATTTA	2827
Db	2659	CAACCTAAAGCCAGAAACCTTCGAGAGTATTTGTTGGGGTGTGTTTTATATTTA	2714
RESULT 6			
S81243			
LOCUS			
DEFINITION			
S81243			
CHN=steroid/thyroid orphan receptor homolog gene [human, fetal			
brain, mRNA Partial, 2714 nt].			
S81243			
VERSION			
S81243.1			
GI:1311504			
KEYWORDS			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata, Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2714)  
Clark, J., Benjamin, H., Gill, S., Sidhar, S., Goodwin, G., Crew, J., Gusterson, B.A., Shipley, J. and Cooper, C.S.  
Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma  
Oncogene 12 (2), 229-235 (1996)  
96152889  
8570200  
GenBank staff at the National Library of Medicine created this entry [NCBI Gibbsq 176289] from the original journal article.  
Location/Qualifiers  
1..2714  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2714  
/\_gene="CHN"  
/note="steroid/thyroid orphan receptor homolog gene; rat NOR1 homolog"  
443..2497  
/\_gene="CHN"  
/note="steroid/thyroid orphan receptor homolog gene; No start codon was found. N-terminal of protein truncated by chromosomal breakpoint 2 nucleotides 5' of first (Pro) codon"  
/codon\_start=1  
/protein\_id="AAB36006.1"  
/db\_xref="GI:1311505"  
/translation="PTAETGSSPASPCPEPGPLAVPGSTAGSPRTSAPPTLSASAC  
TSPFTQARYPDMPQCVOAQVSPSPGSSYAAQYSEYTEINNPDTLKLTDMLD  
TEHTATTSLSPISITFVEGYSNNELPSCVYQMRPLIKVEEGAEPSSHHHHHH  
HHHHHHQHQHQPSIPASSPEDELPSTSMYFKQSPSTPTTAPFPQACALWDE  
PSAPGATGCLLDPPMAQVPTVAGARFLFHFKESPHPAPSPAGPHILGYPGT  
AALLPLGAAAGQAALLESHPYGLPLAKRAAPLAPPLGLTPSPATSSILGES  
LPSPSSSSSGEGTCVAGCDNAACHVYRTCEGCKGFFKRTVQKNAKYCLANKI  
PVDKRRNRQVCRFGKCLSLSVKMVEVTRTDSLKRGRGLFSKPKSPQQPSPQSP  
SPDCKNMLVRALITDPRDLDYKRYCTDQAAGTDAEHVQVFNLLTASIDVSV  
WASKIPGFTDLPKEDQTLTLESFLEFLVRLSIRSNATDAEDKPFVCGNLVLRHQI  
GFEWGLDLSKDFSLNLDIQALACLALSMTTERRHGLKPKRVEELCNKITS  
KHQSKGQALEPTESKLGALVELRKKITGLQIRIFYLKLDLVSPSPSIIIDLFLDY  
PF"

## ORIGIN

	Query Match	70.9%;	Score 2689.6;	DB 9;	Length 2714;	
	Best Local Similarity	99.9%;	Pred. No. 0;			
	Matches 2692;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;	
Qy	132	TCACCTCCACACACAGAGCGCGCACACAGGCTCCGACACACACTTCGCTCTCCC	191			
Db	19	TCACACAAACACACAGACACAAGCGCGCACACAGGCTCCGACACACACTTCGCTCTCCC	78			
Qy	192	CGCGCTCACACCCCTCTTCGCTGAGCCCTTGCCTGCTGACGCGCGCGAGCTGGA	251			
Db	79	CGCGCTCACACCCCTCTTCGCTGAGCCCTTGCCTGCTGACGCGCGCGAGCTGGA	138			
Qy	252	CGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGCGCAGTGGCGCTGGAGGTGGGA	311			
Db	139	CGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGCGCAGTGGCGCTGGAGGTGGGA	198			
Qy	312	ACAGCGCGCATCTCTCCCTCTGGTACACAGCCCAAGCCAGGACGCGCGCGGAACCTCTC	371			
Db	199	ACAGCGCGCATCTCTCCCTCTGGTACACAGCCCAAGCCAGGACGCGCGCGGAACCTCTC	258			
Qy	372	GGCTGTCTCTCCATAGTGGGATCGAGCATCCCCCAACAGCGCTCACCGCTCCG	431			
Db	259	GGCTGTCTCTCCATAGTGGGATCGAGCATCCCCCAACAGCGCTCACCGCTCCG	318			
Qy	432	GGAGCGCTGGGCTTGACACCGCAGCCCTTCGGGGACAGCTGTACTCCCCCCCAG	491			
Db	319	GGAGCGCTGGGCTTGACACCGCAGCCCTTCGGGGACAGCTGTACTCCCCCCCAG	378			





224 GCGGTT-----GCAGCGCGCGCGCAGCTGAGCGCCCTCCGCGGCTCACTTGCACCG 278  
181 GCGGTTGTCGCGCGCGCGCGCGCGCAGCGCGCGCCCTCCAGGGCTCACTTTGCAAG 240  
279 CTGACGGTCCGCGCAGTGGCGGTGGAGGTGGAAACAGCGCGCGCGCATCTCCCGCCCTGGTTC 338  
241 CTGACAGAGCGGCGAGTGGCGGTGGAGGTGGAAACGTGGCGACATCTTAGCCCTTGCTC 300  
339 ACAGCCCAAGCAGAGCGCGCGGAAACCTCTCGGCTGTGCTCTCCCATGATCGGGATC 398  
301 GCAGCGGAGACTGGAGC--CTGCGGAACCTCTCGCGCGCGCTCTCCCATGATGGGATC 359  
399 GCAGCATCCCCACACAGCC--GCTCACCGCTCTCGGGAGCGCTGGGGCTGTACACGC 455  
360 GCAGCATCCCCACAGCGCGCTCTCACCGCTCTGGGAGCGCTGGGGTTGTACACGC 419  
456 AGCCCTTCGCGGACAGCAGCTGTGACTCCCCCGCCAGTGCAGATTTTCGGGACAGCTCTCTA 515  
420 AGCCCTTCGCGGACAGCAGCTGTGACTCTCCCCCAATCCAGATTTTCGGGGTGCCTCTCTA 479  
516 GAAACTCGCTCTAAGACGGAAACCGCCACAGCACTCAAGGCCCATGCGGGAAGAGGGCAG 575  
480 GAAACTCGCTCTAAGACGGAAACCTCCACAGAACCCAAAGCCCATGCGGGAGAGCGCAG 539  
576 CCGCGCAAGCCCGCGGCTGAGCTGGACCTTAGCGGTGCCGGCAGCA-----CTGC 629  
540 CCGCGCAAGCCCGCGGCTGAGCTGGACCTCAACAGAGCGGGCCAGCACAGCGCGCGC 599  
630 CCGCGCTTCGCGTCCGCGGACGTCGCTCTCTCACTCTCAGCTCTCGCTGGAGAGAC 689  
600 GCTGCTTCGCTATCCCGAGCTCCCGGCTCTCTACACTCTCAGCTCTCGCTGGAGAGAC 659  
690 CCGCGAGCCCACTACAGCGCGCAAGATACCTCCAGATATGCCCTGGCTCCAGGCCA 749  
660 CCGCGAGCCCACTACAGCGCGCAAGATACCTCCAGATATGCCCTGGCTCCAGGCCA 719  
750 ATATAGCCCTCCCTCCAGGTTCCAGTTATGGGGCAGACATACAGCTCCGGAATAC 809  
720 ATATAGCCCTTCGCGTCCGCGGTTCACCTTATGCCAGCGAGACTATGGCTCGGAATAC 779  
810 CACGAGATCATGAACCCCGACTACACCAAGCTGACCATGAGCTTGGCAGCACTGAGAT 869  
780 CACAGNAATCATGAACCCCGACTATGCCAAGCTGACCATGAGCTTGGTAGCAGGGAT 839  
870 CACGGCTACAGCCACCATCTCCCTGCCAGCATCAGTACCTTCTGGAGGGTACTTCGAG 929  
840 CATGGCCACCGCCACGAGCTCCCTGCCAGCTTCAGTACCTTCATGGAGGGCTACCCAG 899  
930 GAACCTACGAACTCAAGCTTCCTCGGTGTACCAATGCGAG-----CGGCCCTT 977  
900 CAGCTGCGAACTCAAGCCCTCTGCTGTACCAATGCGCGCTTCTGGGCGCTTCGCGCTTT 959  
978 GATCAAGTGGAGGGCGCGCGCGCTACCATCACCATCACCACACACACCA 1037  
960 GATCAAGATGGAAGGGTTCGGAGCATGCTTACCAACACACACACCATCATCA 1019  
1038 CCACCAACCATCATCAGCAGCAGCATCAGCAGCCATCCATTCCTCCAGCTTCCAGGCC 1097  
1020 TCATCACCACACACACAGCAG-----CAGCAGCGCTCCATTCCTCTCTCTGGCC 1073  
1098 GGAGGACGAGGTGCTGCCAGCACTCCATGTACTTCAAGCAGTCCCAACCGTCCACCC 1157  
1074 CGAGGACGAGGTACTGCCAGCACTCCATGTACTTCAAGCAGTCTCCGCGCTTACGCC 1133  
1158 CACCAAGCGGCTTCCCGCGCAGCGGGGGCTTATGGGACGAGGCACTGCGCTCGGC 1217  
1134 GACCACTCAGGCTTCCCGCGCAGCGGGGGCTGTGGGAGCAGCAGCTGCGCTCTGC 1193  
1218 GCGCGCTCATCGCAACCGCGCGCTGCTGGACCGCGCGATGAAGGGGCTCCCAACGCT 1277  
1194 GCTTGGCTCATCGCTCCGGGACCGCTGTGGACCGCGCAGATGAAGGCAAGTCCCCCAAT 1253  
1278 GCGCGCGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCGCGCGCGCATCCCCCGCGCC 1337

1254 GCGCGCTGCTGCGCGCTTCCCGATCTTCTTAAGCCCTCACGCGCACACCTCTCCCGGCC 1313  
1338 CAGCCCGCGCGCGCGCCACCACTCTGGCTACGACCGAGCGCGCTGCGCGCTCAGCCT 1397  
1314 CAGCCACGCGCGCGCGCCACCACTCTGGCTATGACCCACGCGCGCAGCTCGCTCAGTCT 1373  
1398 GCGCGTGGGAGCGCAGCGCGCGCGCGCAGCGCGCGCGCTTGAAGAGCACCGCTA 1457  
1374 ACCCTTGGAGCGCGCGCGCGCGCGCAGCAAGCTGTGCTGCTGAGGGCGCATCCGTA 1433  
1458 CCGGCTGCGCTGGCGCAAGAGCGCGCGCTGCGCTTCCCGCTCTCGGCTCACGCG 1517  
1434 CCGGCTTCCGCTGGCGCAAGAGCGCGCGCTGAGCTTCCCTCGCTGGGCTCACAGC 1493  
1518 CTCCCTTACCCTCCAGCTCTGGCGGAGAGTCCCAGCTTCCCGTCCCGCGCCAGCAG 1577  
1494 GTCCCTTACCGCTCCAGCTCTGGAGAGAGCGCCAGCTTCCATCGCCACCCCAATAG 1553  
1578 GAGCTGCTGCTTGGCGAGGCGACGCTGCGCTGTGCGGGGACACCGCGCTGCCAGCA 1637  
1554 GAGCTCATCATCCGCGAGGCGACGCTGTGCTGTGCGGGGCAATGTCTGCTGCCAGCA 1613  
1638 CTACCGCGTGGCAACTCTGCGAGGCTGCAAGGCTTTTTCAAGAGAACAGTGCAGAAAA 1697  
1614 CTACGAGTCCGCACTGCGAGGCTGCAAGGCTTCTTCAAGAGAACCGTGCAGAAAA 1673  
1698 TGCAAAATATGTTTCCCTGGCAATAAATACTGCCAGTAGACAGAGACGTCGAAACCG 1757  
1674 CGCAAAATATGTTTCTTGGCAATAAATACTGCCCGTAGACAGAGACGTCGAAATCG 1733  
1758 ATGTCAGTACTGTCGATTTTCAAGAGTCTCAGTGTGGAATGGTAAAGAGATGTCG 1817  
1734 ATGTCAGTACTGTCGATTTTCAAGAGTCTCAGTGTGCGGATGGTGAAGAGATGTCG 1793  
1818 TACAGATAGTCTGAAGGAGGAGAGTCTGCTGCTTCCAAACCAAGAGGCCATTACA 1877  
1794 TACAGATAGTCTGAAGGAGGAGAGTCTGCTGCTTCCAAACCAAGAGGCCATTACA 1853  
1878 ACAGGAACCTTCTCAGCCCTTCCACCTTCTCCTCCAATCTGCATGATGAATGCCCTGT 1937  
1854 ACAGAGCCCTCGCAGCCCTCCCGCCATCTCCTCGATCTGTATGATGAACGCCCTGT 1913  
1938 CCGAGCTTTAAGACACTCAACACCGAGAGATCTGATTTATCCAGATCTGTGCCACTGA 1997  
1914 CCGAGCTTTAAGACAGCAACCGCCAGAGACCTTGAATTACTCCAGATCTGTCCCA 1973  
1998 CCAGGCTGCTGCGCAGCAGATGCTGAGCATGTCGCAAACTTCAACCTCTCTGACAGC 2057  
1974 CAGGCGCTTGGCGCAGACGCTGAGCAGCTGAGCAGCTTCAACCTTCTGACGCGC 2033  
2058 CTCCATTTGATGTATCCAGAGCTGGCGAGAAAGATTCGCGGATTTTACTGATCTCCCAA 2117  
2034 CTCCATCGAGCTTCCAGAGCTGGCGAGAAAGATCCCGGATTCACCTGATCTCCCAA 2093  
2118 AGAAGATCAGACATTAATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTC 2177  
2094 AGAAGATCAGACGTTTACTTATAGAATCAGCCTTTTGGAGCTGTTTGTCTTAGACTTTC 2153  
2178 CATCAGGTCAACACTGCTGGAAGTAAAGTGTGTTCTGCAATGGAATGCTCTGATCG 2237  
2154 TATCAGGTCAACACTGCTGGAAGTAAAGTGTGTTCTGCAATGGAATGCTCTGATCG 2213  
2238 ACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAA 2297  
2214 ACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAA 2273  
2298 TTTGAGAGCTGAACCTTGAATCAAGCCTTAGCCTGCTGTCAGCACTGAGCATGAT 2357  
2274 TTTGAGAGCTGAACCTTGAATCAAGCCTTAGCCTGCTGTCAGCACTGAGTATGAT 2333  
2358 CACAGAAAGACATGGTTTAAAGAACCAAGAGAGTTCGAGAGCTATGCAAGATCAC 2417



Qy	399	GCAGCATCTCCCCACACAGCC---GCTCACCGGCTCTCGGGAGCGCTGGGCTTGTATACCGC	455
Db	360	GCAGCATCTCCCCAGCAGCAGCGCTGCTCTCAACCGCTCTGGGAGCGCTGGGTTGTGACCGC	419
Qy	456	AGCCCTTTCGGGACAGCAGCTGTGCATCTCCCCCAGTGCAGATTTCCGGACAGCTCTCTA	515
Db	420	AGCCCTTTCGGGACAGCAGCTGTGCATCTCCCCCAATCCAGATTTCCGGGTGCTCTCTA	479
Qy	516	GAATACTCGCTCTAAAGACGGAACCGCCACAGCACTCAAAGCCCACTCGGAAGAGGCGAG	575
Db	480	GAATACTCGCTCTAAAGACGGAACTCTCACAGAACCCAAAGCCCACTCGGGAGAGCGCAG	539
Qy	576	CCGGGCAAGCCCGGGCCCTGAGCCCTGACCCCTTAGCCGTTAGCCGTCCGGGGACGA	629
Db	540	CCCGACAAGCCCGGGCGTGAAGCTGGAACCCCTCAACAGAGCGGCCACAGCAGCGCGGC	599
Qy	630	CGGCGCTTTCGCTCTCGCCGAGCTCCGCTCTCTCTTACACTCTTCAGCCCTCCGCTGGAGAGAC	689
Db	600	GGCTGCTTTCGCTATCTCCGAGCTCCCGCTCTCTTACACTCTTCAGCCCTCCGCTGGAGAGAC	659
Qy	690	CCCCAGCCCCACCAATTGAGCGCGCAAGATACCCCTCCAGATATGCCCTTCGCTCAAAGCCCA	749
Db	660	CCCCAGCCCCACCAATTGAGCGCGCAAGATACCCCTCCAGATATGCCCTTCGCTCAAAGCCCA	719
Qy	750	ATATAGGCTTCCCTTCAGGTTCCAGTTATGCGGCGCAGACATACAGCTTCGGAAATACAC	809
Db	720	ATATAGGCTTTCGCTCTCGGGGTCCACTTATGCCACGCGAGACTTATGGCTCGGAAATACAC	779
Qy	810	CACGGAGATCATGAACCCCGACTACACCAAGCTCACCATGACCATCTGGCAGCACTGAGAT	869
Db	780	CACAGAAATCATGAACCCCGACTATGCAAGCTGACCATGACACTCGTAGCAGCGGGAT	839
Qy	870	CAGGCTACAGCCACCAAGTCCCTGCCAGCATCAGTACCTTCGTGAGGGCTACTCGAG	929
Db	840	CATGGCCACGSCCAGCAGCTCCCTGCCAGCTTCAGTACCTTCATGAGGGCTACCCCGAG	899
Qy	930	CNACTAGCACTCAAGCTTTCCTGGGTGTACCAATGCAG-----CGGCCCTT	977
Db	900	CAGCTGGAATCAAGGCCCTCTGCTGTGTACCAATGCCGCTTCTGGGCCCTCGGCCCTT	959
Qy	978	GATCAAAAGTGGAGAGGGGGCGGCCAGCTACCATCACCATCACCAACACCAACCA	1037
Db	960	GATCAAGATGGAAGAGGGTCCGAGCATGGCTACCAACAACCAACCATCACCATCA	1019
Qy	1038	CCACCACCAATCAACAGCAGCAGCATCAGAGCCATTCATTCCTTCAGGCTCCAGGCC	1097
Db	1020	TCATCACCAACCAACAGCAG-----CAGCAGCGCTCCATTCCTCTCCCTCTGGGCC	1073
Qy	1098	GGAGGACGAGGTGTGCCCGCAGCACTCCATGTACTTCACGAGTCCCAACGTCACCCC	1157
Db	1074	CGAGGACGAGGATCTGCCCGCAGCACTCCATGTACTTCACGAGTCTCCGCGCTCTAGGCC	1133
Qy	1158	CACCACCGCGCTTTCCTCCCGCAGGCGGGGGCTTATGGGACGAGGCACTGCCCTCGGC	1217
Db	1134	GACCACTCCAGCTTTCCTCCCGCAGCGGGGGCGCTGTGGGACGACGAGTGCCTCTGCG	1193
Qy	1218	GCCCGGCTGCATCGCAACCGGCCCGCTGTGTGAACCCGCGATGAAGGGGGTTCGCCACGGT	1277
Db	1194	GCCTGGGTGCATCGCTCTCGGAGCCGCTGTGGAACCCCGCAGATGAAGGCACTGTCGCCCAAT	1253
Qy	1278	GGCCGGCGCGGCTTTCGCGTCTTTCATCTTCAAGCCCTCCGCGCGCATCCCCCGGGCC	1337
Db	1254	GGCCGCTGTGCGGCTTTCGGAATCTTCTTCAAGCCCTCACCGCCACACCTTCCCGGGCC	1313
Qy	1338	CAGCCCGCGCGGGCCACCACTTCGGCTACGAACCCGACGCGCCCTGCGCGCTCAGCGCT	1397
Db	1314	CAGCCACGCGCGGCCACCACTTGGCTATGACCCCAACGCGCAGCTGGGCTCAGTCT	1373
Qy	1398	GCGCTGGAGCCGACGCGCGCGGGAGCCAGGGCCGCGCGCTGTGAGAGCCACCGGTA	1457
Db	1374	ACCCCTGGAGCCGCGCGCGGGCAGCAAGCTGTGCTGCTCGAGGAGCATTCGGTA	1433

Qy	1458	CGGGCTCCGGTGGCCAAAGAGGGCGGGCCCGCTGGCCGCTTCCCGGCTCTCGGCTCTCAGCC	1511
Db	1434	CGGGCTCCCGTGGCCAAAGAGGACGGCACGTTGACCTTCCCTCCGCTGGGCGCTCACAGC	1493
Qy	1518	CTCCCTTACCAGCTCCAGCCTGCTTGGCGAGAGTCCAGAGCTGCGCGTCCGCGCCACAGCAG	1577
Db	1494	GTCCCTTACCAGCTCCAGCCTGCTTGGGAGAGGCCCGAGCTTACCATCGCCACCCANTAG	1553
Qy	1578	GAGCTCGTCTGGCGAGGGCCAGTGTGCGCTGTGCGGGGACAAAGCCGCTGCCAGCA	1637
Db	1554	GAGCTCATCATCGCGCGAGGGCCAGTGTGCTGTGTCGGGAGCAATGCTGCTGCTGCCAGCA	1613
Qy	1638	CTACGGCGTGGGAACCTTCGAGGGCTGCCAAGGGCTTTTTCAGAGAAACAGTCGAGAAAAA	1697
Db	1614	CTACGGAGTCGCAACCTTCGAGGGCTGCCAAGGGCTTCTTCAGAGAAACGGTCGAGAAAAA	1673
Qy	1698	TGCAAAATATGTTTTCGCTGCGCAATAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG	1757
Db	1674	CGCAAAATATGTTTTCGCTTGGCAATAAAACTGCGCGGTAGACAAGAGAGCTCGAAATCG	1733
Qy	1758	ATGTCAGTACTGTTCGATTTTCAGAAAGTGTCTCAGTGTGTTGGAATGGTAAAAAGAGTTGTCCG	1817
Db	1734	ATGTCAGTACTGCAGGTTTCAGAAAGTGTCTCAGTGTGCGGATGGTGAAGAGTGTGTGCG	1793
Qy	1818	TACAGATAGTCTGAAGGGAGGAGAGTGCTGTGCTTCCAAACCAAGAGCCCATTTACA	1877
Db	1794	TACAGATAGTCTGAAGGGAGGAGAGTGCTGTGCTTCCAAACCAAGAGCCCACTACA	1853
Qy	1878	ACAGGAACCTTCTCAGGCCCTCTCCACCTTCTCCTCCAAATCTGCATGATGATGCCCTGT	1937
Db	1854	ACAGGAGCCCTTCGAGGCCCTCCCAACCATCTCTCCGATCTGTATGATGAACGCCCTGT	1913
Qy	1938	CCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTATTCAGATACTGTGCCACTGA	1997
Db	1914	CCGAGCTTTAACAGACGCAACGCCAGAGACCTTGATTACTCCAGATACTGTGCCACGA	1973
Qy	1998	CCAGGCTGCTGCAGGCAAGATGCTGAGCATGTGCACAAATTTACAACTCTCTGACAGC	2057
Db	1974	CCAGGCCACTGCGGGCAGACAGCGCTGAGCAGTGCAGCAGTTCTACAACCTTCTGACGGC	2033
Qy	2058	CTCCATTGATGTATCCAGAGCTGGGCGAGAAAGATTCGGGATTTACTGATCTCCCAA	2117
Db	2034	CTCCATCGAGCTGTCCAGAGCTGGGCGAGAAAGATCCCGGATTCAGTATCTCCCAA	2093
Qy	2118	AGAGATCAGACATTACTTATGAACTCAGCGCTTTTGGAGCTGTTGTCTCCAGACTTTC	2177
Db	2094	AGAGATCAGAGCTTACTTATAGAACTAGCGCTTTTGGAGCTGTTGTTCTTAGACTTTC	2153
Qy	2178	CATCAGTCAACAACCTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCAATCG	2237
Db	2154	TATCAGGTCAAAACACTGCTGAAGATAGTTTGTGTTCTGCAATGGACTTGTCTGCAACCG	2213
Qy	2238	ACTTCAGTGCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCCTTAAA	2297
Db	2214	ACTTCAGTGCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTTAAA	2273
Qy	2298	TTTGCAGAGCTGAACCTTGATATCCAAGCTTTAGCTGCTGCTGCAGCTCAGCATGAT	2357
Db	2274	TTTGCAGAGCTGAACCTTGATATCCAAAGCTTTAGCTGCTGCTGCAGCTCAGCATGAT	2333
Qy	2358	CACAGAAAGACATGGGTTAAAGAACCAAAAGAGAGTCGAAGACTATGCAACAAAGATCAC	2417
Db	2334	CACAGCGGACATGGGTTAAAGAACCAAAAGAGAGTGGAGGAGCTATGCAACAAAGATCAC	2393
Qy	2418	AAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGGCCCAACCGAGTCCAGGT	2477
Db	2394	AAGCAGCTTAAAGGACCAACAGAGGAGGACAGGCTCTGGAGGCCCTCAGAGCCCAAGGT	2453
Qy	2478	CCTGGGTGCCCTGGTAGAACTGAGGAGAGATCTGCACCTCGGGCTCCAGCGCATCTTCTA	2537
Db	2454	CCTTCCGCGCATGGTGGAACTGAGGAGAAATCTGCACCCAGGGCTCCAGCGTATCTTCTA	2513
Qy	2538	CCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTGACAAGCTCTTCTCGGACAC	2597

Db 2514 CCTGAAGCTGGAGGACTTGGTGTCCCACTTCTGTCTATCGACAAGCTCTCTCTTGATAC 2573  
 Qy 2598 CTTACTCTTCTAATCAGGAGCAG-TGGAGCAGTGAGCTGCTCTCTCTCTAGCAGCTGCT 2656  
 Db 2574 CTTGCTCTTCTGAGCAGGAGGAGCCTGAGCAGAGAGCTACTTGTCTCTGCTGCACTGTC 2633  
 Qy 2657 TGTACGACGAGAAAGGATAGGTTTGGAAACCTATCATTTCTGCTGCTCTCTTCTTAAGAGGA 2716  
 Db 2634 ATTAAGTGAGCAAAAGGATGGTTTGAACACCT-GCCCTCTATCTCTCTCCAGGGAA 2692  
 Qy 2717 AAAGCAGCTCTGTAGAAAGCAAAAGACCTTTCTTTTCTGCTGCTCTTTCTCTTCAACCT 2776  
 Db 2693 AAAGCAGCTCCCATAGAAAGCAAGACCTT-TTTTCTCTGSCACCTTTCTCTTCAACCT 2751  
 Qy 2777 TAAAGCCAGAAACTTGCAGAGTATTGTGTGGGTGTGGTTTATATATTAGGCAATGGG 2836  
 Db 2752 TAAAGCCAGAACTTGCAGAGTATTGTGTGGGTGTGGTTTATATATTAGGCTTTGT 2811  
 Qy 2837 GGATGGGTGGAGGGGT--TATAGTTTCATGAGGGTTTCTTAAGAAATTTGCTAAACAAG 2894  
 Db 2812 GGTGGGTGGAGGGGTAAATAGTTTCATGAGGCTTTCTTAAGAAATTTGCTAGCAAG 2871  
 Qy 2895 CACTTTTGGACAATGCTATCCAGCAG-----GAAAAAAGGATATATTAACGTGTTT 2948  
 Db 2872 CACTTTTGGATGATCTATCCAGCAGTGGGTGGGAGAAAGGATATATTAACGTGTTT 2931  
 Qy 2949 AAAA--CTCTTCTGGGAATCCAATTATAGTTGCTTTGTTATTTAAACCAAGAACAGCC 3006  
 Db 2932 AAAAATCTTCCGGGGGATATGACTAGTTGCTTTGTTATTTAAATTAAGAACAGCC 2991  
 Qy 3007 AAGGGTTC-TTCCGACGGTAGGATGTGTCTTAAAGATTGGTCTTGAATAATGCTTC 3065  
 Db 2992 AAGGCTGTTTTACCAGGATGGGTGTCTTAAAGACTGATCCCTTTAGTATGATCTC 3051  
 Qy 3066 CTGTATCAAGTAGTATGTGGTGCACCAAGAGCAGAAA-----CTTCCTTT 3113  
 Db 3052 CCGGATC-GAGGCACATAAGTGGTGCAATAGAGCGGGGAAATCTTCATTTCTTCATTT 3110  
 Qy 3114 TAAATTCCTTCTCTTATTTTAAATAATGATGAGGATGAGGATACCTTACAAATCA 3173  
 Db 3111 CTTCCTCTCTTAAATAATGATGAGGATGAGGATGAGGATATCTACAAATCA 3170  
 Qy 3174 GACATGGCAAAATAATAGCTGTGTTGCTTCCATAAACAAGTGCATTTTAAAGTGC 3233  
 Db 3171 GACTTAGCAAAATGATATAGGCTATTCGCTTCCATACAGTGCATTTTATAGAGTGC 3230  
 Qy 3234 TGTCTTACTAGTCTGTTTATTAATCTCTCTTATCTATATGAAATATAAAGAGGC 3293  
 Db 3231 TGTCTTACTAGTCTGTTTGTGAACTCTCCCTCATTTTATATGAAATAAAGAGAGGC 3290  
 Qy 3294 AGTCATGTTAGCAATGACAGCTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGT 3353  
 Db 3291 AGTCATGTTATCAACAGGAGTGTCTATTTCTTACTCTACCTTGGTCCACCTGCTGT 3350  
 Qy 3354 CGATCCCTTCTGAGGTATGGCCATCCAGAGCTTTTAGGCCATTTCTGATGGAACAGAT 3413  
 Db 3351 AGAACCTTCTGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGATGAAATTCGAC 3410  
 Qy 3414 CCC-----TGCCCTGACTGCTCCAGCTATCTCTGAAAGTGGATCAGATTATAAAGT 3463  
 Db 3411 ACCCTCCCTTCAACCCATGACTATCCAGATGCTCTGAAAGTGGGATCAGGTTATAAATG 3470  
 Qy 3464 GATTACATGTAAGTCTGTTTGTGTTGTGT-TCTATCAACCCACAGAGTTCCTTAACTTG 3522  
 Db 3471 GATTGCATAGTCTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3530  
 Qy 3523 CTTAGTTATAGTAAGTCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3575  
 Db 3531 CTTTGTGTGTAGCAAGTTCCTGATTCCTCCATTCAGAGCCCAAGGAGCATTGGGTGACT 3590  
 Qy 3576 TGAGTATTGATCCCTAGATAGACATGCAATCAGCAGGACT-GGTCTACACAGGTA 3634

Db 3591 CGATCAAGGGTTAACCCCTAGGAGAACATGCAAAATAAGTAGGAACTGGGTTCAGACAGGTA 3650  
 Qy 3635 AGCACCAGGCAACAATAAGGATTTTATAGATAT-----AATTTAATTTTGTATTGG 3687  
 Db 3651 AGCACCAGATGATAGGATTTATATATAATATATATAATAATTTTGTATTGG 3710  
 Qy 3688 TTAAGGACAAATTTTGGAGCAAGCAAAATCTTTT 3723  
 Db 3711 TTA--TAGACAATTTTGGAAAGCAAGAGATCATCT 3744  
 RESULT 9  
 RATNORI  
 LOCUS  
 DEFINITION  
 Rattus norvegicus mRNA for NOR-1, complete cds.  
 ACCESSION  
 D38530  
 VERSION  
 D38530.1 GI:1483194  
 KEYWORDS  
 NOR-1; neuron derived orphan receptor; thyroid/steroid receptor superfamily.  
 SOURCE  
 Rattus norvegicus (Norway rat)  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE  
 1 (sites)  
 AUTHORS  
 Ohkura,N., Hijikuro,M., Yamamoto,A. and Miki,K.  
 TITLE  
 Molecular cloning of a novel thyroid/steroid receptor superfamily gene from cultured rat neuronal cells  
 JOURNAL  
 Biochem. Biophys. Res. Commun. 205 (3), 1959-1965 (1994)  
 MEDLINE  
 95110348  
 PUBMED  
 7811288  
 REFERENCE  
 2 (bases 1 to 4400)  
 AUTHORS  
 Ohkura,N.  
 JOURNAL  
 Unpublished  
 REFERENCE  
 3 (bases 1 to 4400)  
 AUTHORS  
 Ohkura,N.  
 JOURNAL  
 Direct Submission  
 TITLE  
 Submitted (14-OCT-1994) Naganari Ohkura, National Cancer Center Research Institute, Growth Factor Division; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan (E-mail:nohkura@ncc.go.jp, Tel:81-3-3542-2511(ex.4302), Fax:81-3-3542-8170)  
 COMMENT  
 On Aug 8, 1996 this sequence version replaced gi:643599.  
 FEATURES  
 Location/Qualifiers  
 1..4400  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Wistar"  
 /db\_xref="taxon:10116"  
 700..2586  
 /note="thyroid/steroid receptor superfamily"  
 /codon\_start=1  
 /product="NOR-1"  
 /protein\_id="BAA07535.1"  
 /db\_xref="GI:643600"  
 /translation="MPCVQAQSPSPGSGTATQYGYEYTEINPNVAKLTMDLGS TGMATATSLPFSFSTMEGYPSCLQYMPSPGRLIKMEGRGHEGYHHH HHHHHHHHQOQSPISPPSGDEVLPSTSMYFKQSPSTPTTGGPPQAGALWDD ELPSAGCIIAGPLDDPMKAVPPMAAARFPIFKPSPHPHPAPSPAGHHGLGYDT AAALSLPLGAAAGSQAALGHPYGLPLAKETATLTFFPLGLTASPTASSLLGES PSLSPNRRSSSGEGTCVCGDNNAACHYGVRCGCKGFFKRTVQKNKXVCLANK NCPVDRNRRCQYCFKQCLUSGVMEVVRDLSLKGRRRLPSKPSLPQQEPSPQS PPSPIICMMNALVRALTADPRDLDSRYCPTDQATAGTDAHVQOQFYNLLTASDV SRSWAKIPGFTDLKPEDOTLLIESAFLEFLVRLISRSNTAEDKFEVFCNGLVHRLQ LRGFGWLDLSIKDFSLNLSLNDIOALACLALSMTIERHGLKPKVEELCNKITS SLKDHQRKQALEPSEPKVLRLVELRKLCTQGLQRIFYKLLEDLVSPSPVLDKFLD TLPP"  
 2587..>4400  
 polyA\_signal  
 3126..3131  
 polyA\_signal  
 4115..4120  
 ORIGIN  
 3'UTR  
 2587..>4400  
 polyA\_signal  
 3126..3131  
 polyA\_signal  
 4115..4120  
 ORIGIN  
 Query Match 65.9%; Score 2500; DB 10; Length 4400;  
 Best Local Similarity 83.4%; Pred. No. 0;



QY 2178 CATCAGGTCAACACATGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTTGTCTCGATCG 2237  
DB 2154 TATCAGGTCAACACATGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTTGTCTCGATCG 2213  
QY 2238 ACTTCAGTGCCTTCGTGGATTGGGAGTGGCTCGACTCTATTAAAGACTTTTCCCTTAA 2297  
DB 2214 ACTTCAGTGCCTTCGTGGATTGGGAGTGGCTCGACTCTATTAAAGACTTTTCCCTTAA 2273  
QY 2298 TTTTGAGAGCCTGAACCTTTGATATCCAAGCCTTAGCTGCTGTCAGCACTGAGCATGAT 2357  
DB 2274 TTTTGAGAGCCTGAACCTTTGATATCCAAGCCTTAGCTGCTGTCAGCACTGAGCATGAT 2333  
QY 2358 CACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAAGATCAC 2417  
DB 2334 CACAGAGCAGCATGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAAGATCAC 2393  
QY 2418 AAGCAGTTTAAAGAACCAAGAGTGAAGAGAGCTGCGAGCCACCGAGTCCAAGGT 2477  
DB 2394 AAGCAGCTTAAAGAACCAAGAGAGAGGAGCAGGCTCTGGAGCCCTCAGAGCCCAAGGT 2453  
QY 2478 CCGTGGTGCCTTGGTAGAAGTCTGCAACCTTGGCCCTCCAGCGCATCTTCTA 2537  
DB 2454 CTTTCGCGACCTTGGTAGAAGTCTGCAACCTTGGCCCTCCAGCGCATCTTCTA 2513  
QY 2538 CTTGAAGCTGGAAGACTTGGTCTCCACCTTCCATCAATGCAAGCTCTTCCCTGACAC 2597  
DB 2514 CTTGAAGCTGGAAGACTTGGTCTCCACCTTCTGTCATCGAACAAGCTCTTCCCTGATAC 2573  
QY 2598 CTTACTCTTCAATCAGGAGCAG-TGGAGCAGTGAAGCTGCTCCCTCTCCTAGCACCTGCT 2656  
DB 2574 CTTGCTTCTCAGCAGGGAAGCTGAGCAGAGACTACTTGTCTGCTGCACTGGTCTC 2633  
QY 2657 TGTCTAGCAGCAAGGGATAGGTTTGGAAACCTATCATTTCTGCTCCTTCTTAAGAGGA 2716  
DB 2634 ATTAAGTGAACAAAGAGTGGTTTGAACCT-CCCTCTATCTCTCCTCCAGGGGA 2692  
QY 2717 AAAGCAGCTCTGTGAAGCAAGAACTTTCTTTTCTGCTGCTCTTCTCCTTCAACACC 2776  
DB 2693 AAAGCAGCTCTCCTGAAGCAAGAACTTT-TTTTTCTGCTGCTCTTCTCCTTCAACACC 2751  
QY 2777 TAAAGCCAGAACTTGCAGAGTATGTTGGGGTGTGTTTATATTTAGGCAATCGG 2836  
DB 2752 TAAAGCCAGAACTTGCAGAGTATGTTGGGGTGTGTTTATATTTAGGCTTTGGT 2811  
QY 2837 GGATGGGTGGAGGGGT--TATAGTTTCAATGAGGTTTCTTAAGAAATTTGCTAACAAAG 2894  
DB 2812 GGTGGCTGGAGGGGTAAATAGTTTCAAGAGCTTTCTTAAGAAATTTGCTAGCAGAG 2871  
QY 2895 CACTTTTGGACATGCTATCCAGCAG-----GAAAAAAGGATATATTAACCTGTTT 2948  
DB 2872 CACTTTTGGATGATCTATCCAGCAGTGGGTGGGGAAGGATATATTAACCTGTTT 2931  
QY 2949 AAAA--CTCTTCTGGGAATCCAAATATAGTTGCTTGTATTTTAAACAGACAGCC 3006  
DB 2932 AAAAATCTTTCCGGGGGAATAGTACTATGTTGCTTGTATTTTAAATTAAGACAGCC 2991  
QY 3007 AAGGGTTG-TTCCAGGAGTGGATGCTGTCTTAAGATTTGGTCCCTTGAATAATGCTTC 3065  
DB 2992 AAGGGCTGTTTACCAGGAGTGGCTGTCTTAAAGATGATCCCTTTAGTATGTACTTC 3051  
QY 3066 CTGTATCAAGGTAGTATGCTGTCGAACCAAGGAGCAAAA-----CTTCCCTTT 3113  
DB 3052 CCGGATC-GAGGCACATAAGTGGTGAATAGAGGGGGGGAATTTCTTCATTTCTTCATTT 3110  
QY 3114 TAAATTCCTCTCTCTTTATTTTAAACAAATGGTGAAGATGAGGATTTACCTACAAATCA 3173  
DB 3111 CTTTCTCTCTCTTAAATTAAGTGAACCAAGGATGAGGATTTATCTACAAATCA 3170  
QY 3174 GACATGGCAAAACAATATGCTGTTTGTCTCCATAAACAAGTCAATTTTAAAGTGC 3233  
DB 3171 GACTTAGCAAAATGATATGGCTATTCGCTTCCACATACAGTCAATTTTATAGTGC 3230  
QY 3234 TGTCTTACTAGTCTTGTATTACTCTCTCTTTTATCTATATGGAATTAAGAGGCG 3293

DB 3231 TGTCTTACTAGTCTTGTGTTGTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3290  
QY 3294 AGTCAATGTTAGCAATGACAGTAAATATCCTACAGAGGCTGTGTTCACTTCCCTTGT 3353  
DB 3291 AGTCAATGTTCAAAAGCGGTGCTCATTTTCTAGCTCACCTTGGTCCACTGCTGCTGT 3350  
QY 3354 CGATCCCTTCTGAGGTATGGCCCATCCAAAGACTTTTATAGGCCATTTCTGATGGAACAGAT 3413  
DB 3351 AGAACCTTTCGAGGTATGGCCCTTCTAAGACTTTTCAGGCCACTCTTGTATGGAATTCGAC 3410  
QY 3414 CCC-----TGCCCTGACTGCTCCAGCTATCTCTGAAAGTGGATCAGATTTATAAATG 3463  
DB 3411 ACCCTCCCTCAACCCATGACTATCCAGATGCTCTGAAATGGGATCAGGTTTATAAATG 3470  
QY 3464 GATTACATGTAACCTGTTTGTGTTGTTGTT-TCTATCAACCCACAGAGTTCCTTAAACTTG 3522  
DB 3471 GATTGATATGACTGTTGTTGCTGTTGTTGTTGTTCACTTGGACAGAGTTCCTTAACTTT 3530  
QY 3523 CTTCAAGTTATAGTAACTGACTGGTATATTTCTTATTCAGAAAGC-----GCCATAAGTCACT 3575  
DB 3531 CTTTATGTTAGCAAGTTCTCTGATTCCTTCAATTCAGAAAGCCCAAGGAGCATTTGGGTGACT 3590  
QY 3576 TGAGTATTTGATCCTTAGATAAGAACTGCAAAATCAGCAGGAAT-GGTATACAGGTA 3634  
DB 3591 CGATCAAGGGTTAAACCTTAGGAGAACATGCAAAATAGTAGGAATCGGTTCAGACAGGTA 3650  
QY 3635 AGCACCAGGGAACAATAAGGATTTTATAGATAT-----AATTTAAATTTTGTATTGG 3687  
DB 3651 AGCACAGAGATGATAGGATTTTATATAATATATAATAAATTAATTTTGTATTGG 3710  
QY 3688 TTAAGGAGCAACTTTTGGAGAGCAAGCAAACTTTT 3723  
DB 3711 TTA--TAGACAATTTTGAAGCAAGAGATCATCT 3744

RESULT 10  
BC068150  
LOCUS  
DEFINITION  
Mus musculus nuclear receptor subfamily 4, group A, member 3, mRNA  
ACCESSION  
BC068150  
VERSION  
BC068150.1 GI:45829690  
KEYWORDS  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 5586)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heise, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M.E., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinska, M.I., Skalska, U., Smal, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 5586)  
Strausberg, R.









[illegible]

RESULT 12					
HSNUCREC					
LOCUS	HSNUCREC	2546 bp	mRNA		
DEFINITION	H.sapiens mRNA for nuclear receptor.				
			linear	PRI 22-JAN-1996	

```

ACCESSION X89894
VERSION 89894.1 GI:1165104
KEYWORDS nuclear receptor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Labelle, Y.; Zucman, J.; Stenman, G.; Kindblom, L.G.; Knight, J.;
Turc-Carel, C.; Dockhorn-Dworniczak, B.; Mandahl, N.; Desmaze, C.;
Peter, M.; Aurias, A.; Delattre, O. and Thomas, G.
TITLE Oncogenic conversion of a novel orphan nuclear receptor by
chromosome translocation
JOURNAL Hum. Mol. Genet. 4 (12), 2219-2226 (1995)
MEDLINE 96177652
PUBMED 8634690
REFERENCE 2 (bases 1 to 2546)
AUTHORS Labelle, Y.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1995) Y. Labelle, INSTITUT CURIE, SECTION DE
RECHERCHE, 26 RUE D'ULM, 75231 PARIS CEDEX 05, FRANCE
FEATURES
    source
        1..2546
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="TEC 1 and TEC 5"
            /tissue_type="heart"
            /clone_lib="lambda GT11"
            /dev_stage="20-26 week old fetus"
            688..2019
            /codon_start=1
            /product="nuclear receptor"
            /protein_id="CAA61984.1"
            /db_xref="GI:1165105"
            /db_xref="GOA:Q14979"
            /db_xref="GOA:Q92570"
            /db_xref="UniProt/Swiss-Prot:Q92570"
            /translation="MPCVQAQYSPGGSSYAAQYSSBYTEINWPNYDKLTMDLGS
            TEITATLSPISITFVEGYSNLYLKPSCYVQWRPLIKVEGRAPSYHHHHHHH
            HHHHHQOHOQSPSPASPEDEVLPTSMYFKQSPSPSTPTTAPFPAQALMDEAL
            PSAPCIAPQLDIPMKAVPTVAGARPLFHFKPSPPHPPAPSPAGGHILGYDPTAA
            AALSLPLGAAAGSQAALSHPYGLPLAKRAAPLAPPLGLTSPSTASSILGSPS
            LPSPSRSSSEGTCAVGDNAACQHYGVRTCEGKGFPRKTVQKNAYVCLANKNC
            PVKRRNRNCQTRFQKLSVMKREVTSLKGRRLPLSPKSPLOQBPSPSP
            SPPICMNALVRALTDSTPRDLIDYSRVSMISCFQMDQGLYLWLLVIRVD"
ORIGIN
    Query Match 51.1%; Score 1939.4; DB 9; Length 2546;
    Best Local Similarity 99.9%; Pred. No. 1.7e-305;
    Matches 1940; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 43 GAGCGAGTCTCTGCTCCCGCCGCCACCCCTCCAGCTCCTGCTCCTCTCCGCTCC 102
DB 1 GAGCGAGTCTCTGCTCCCGCCGCCACCCCTCCAGCTCCTGCTCCTCTCCGCTCC 60
QY 103 CATACACAGCGGTTCACACCGGTCTCTCTCTGTCACACACACACACACAGCGGCACA 162
DB 61 CATACACAGCGGTTCACACCGGTCTCTCTCTGTCACACACACACACAGCGGCACA 120
QY 163 CAGGCTCCGACACACACTTCTCTCCCGCGGTTCACACCCCTTTCCTCCGCTCCT 222
DB 121 CAGGCTCCGACACACACTTCTCTCTCCCGCGGTTCACACCCCTTTCCTCCGCTCCT 180
QY 223 TCCCGGTGACGCGCGCGCGAGCTGGACGCCCTCCCGGTTCACCTTTGCAACGCTGA 282
DB 181 TCCCGGTGACGCGCGCGCGAGCTGGACGCCCTCCCGGTTCACCTTTGCAACGCTGA 240
QY 283 CGGTGCGCGAGTGGCGGTGGAGGTGGAAACAGCGCGGCATCTCTCCCGCTGGTCA 342
DB 241 CGGTGCGCGAGTGGCGGTGGAGGTGGAAACAGCGCGGCATCTCTCCCGCTGGTCA 300
QY 343 CCCAAGCCAGGAGCGCGCGGAAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAG 402

```

```

|||||
301 CCCAAGCCAGGAGCGCGCGGAAACCTCTCGGCTGTGCTCTCCATGAGTCGGGATCGCAG 360
QY
403 CATCCCCACACAGCGCTCACCGCTTCGGGAGCGCGCTTGTGTACACCGAGCCCTT 462
|||||
361 CATCCCCACACAGCGCTCACCGCTTCGGGAGCGCGCTTGTGTACACCGAGCCCTT 420
DB
463 CGGAGACAGAGCTGTGACTCCCGCCAGTGCAGATTTGGGACAGCTCTCTAGAACTC 522
|||||
421 CGGAGACAGAGCTGTGACTCCCGCCAGTGCAGATTTGGGAGAGCTCTCTAGAACTC 480
DB
523 GCTCTAAAGACGGAACCGCCACAGCACTCAAGCCCACTTGCAGAGAGGCGGCGGCA 582
|||||
481 GCTCTAAAGACGGAACCGCCACAGCACTCAAGCCCACTTGCAGAGAGGCGGCGGCA 540
DB
583 AGCCCGGCGCTGAGCGCTTGAGCCCTTAGCGGTGCGGCGAGCACTGCGCGGCTTCGCT 642
|||||
541 AGCCCGGCGCTGAGCGCTTGAGCCCTTAGCGGTGCGGCGAGCACTGCGCGGCTTCGCT 600
DB
643 CGCGGAGCTCGGCTCTCTTACACTCTCAGCTCGGCTGGGAGAGAGCCCGAGCCGAC 702
|||||
601 CGCGGAGCTCGGCTCTCTTACACTCTCAGCTCGGCTGGGAGAGAGCCCGAGCCGAC 660
DB
703 ATTCAGCGCAAGATACCTCCAGATATGCCCTGCTCAAGCCCAATATAGCCCTTCC 762
|||||
661 ATTCAGCGCAAGATATCCCTCAGATATGCCCTGCTCAAGCCCAATATAGCCCTTCC 720
DB
763 CTTCCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATG 822
|||||
721 CTTCCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATG 780
DB
823 AACCCGAGTACACCAAGCTGACCTTGGCAGCACTGAGATCAGGCTCAGGCC 882
|||||
781 AACCCGAGTACACCAAGCTGACCTTGGCAGCACTGAGATCAGGCTCAGGCC 840
DB
883 ACCAGTCTCTGCGCCAGCATCAGTACTTCTGAGGAGGCTACTTCAGAGCAACTACGAACTC 942
|||||
841 ACCAGTCTCTGCGCCAGCATCAGTACTTCTGAGGAGGCTACTTCAGAGCAACTACGAACTC 900
DB
943 AAGCTTCTCTGCTGTACAAATGAGCGGCTTGTATCAAAAGTGAGAGGCGGCGG 1002
DB
901 AAGCTTCTCTGCTGTACAAATGAGCGGCTTGTATCAAAAGTGAGAGGCGGCGG 960
QY
1003 CCCAGTACCATCACCATCACACACACACACACACACACACACACACACACACACACAGCAG 1062
DB
961 CCCAGTACCATCACCATCACACACACACACACACACACACACACACACACACACACAGCAG 1020
QY
1063 CATCAGCAGCATCCATTCCTCCAGCTCCAGCCCGAGGAGGAGTGTGCTGCCAGCACC 1122
DB
1021 CATCAGCAGCATCCATTCCTCCAGCTCCAGCCCGAGGAGGAGTGTGCTGCCAGCACC 1080
QY
1123 TCATGTACTTCAAGCAGTCCCACTCCACCCACACACACCGCGCTTCCCGCCGAG 1182
DB
1081 TCATGTACTTCAAGCAGTCCCACTCCACCCACACACCGCGCTTCCCGCCGAG 1140
QY
1183 GCGGGGCGTTATGGGACGAGCACTGCGCTCGGCGCGCGGTGATCGACCGCGCG 1242
DB
1141 GCGGGGCGTTATGGGACGAGCACTGCGCTCGGCGCGCGGTGATCGACCGCGCG 1200
QY
1243 CTGCTGGACCGCGCATGAAGCGGTGCCACGCTGGCGCGCGGTGATCGACCGCGCTTC 1302
DB
1201 CTGCTGGACCGCGCATGAAGCGGTGCCACGCTGGCGCGCGGTGATCGACCGCGCTTC 1260
QY
1303 CACTTCAAGCCCTCGCGCGCATCCCGCGCGCGAGCCCGCGCGCGGCGGCGGCAACCTTC 1362
DB
1261 CACTTCAAGCCCTCGCGCGCATCCCGCGCGCGAGCCCGCGCGCGGCGGCGGCAACCTTC 1320
QY
1363 GGTACGACCGAGCGCGCTGCGCGGTGAGCTTGCCTGGGAGCGCGGCGCGCGG 1422
DB
1321 GGTACGACCGAGCGCGCTGCGCGGTGAGCTTGCCTGGGAGCGCGGCGCGG 1380
QY
1423 GGCACGAGCGCGCGCTTGGAGAGCCACCGCTAGCGGCTCGCGCTGCGGAGGCGG 1482

```



QY	1927	AATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTATTCAGATAC	1986
Db	1201	AATGCCCTTGTCCGAGCTTTAAACAGATGCAACACCCAGAGATCTTGATTATTCAGATAC	1260
QY	1987	TGTCCACTGACAGGCTGCTGCAGGACACAGATGCTGAGCATGTGCAACAAATTCACAC	2046
Db	1261	TGTCCACCCAGCAGGCCACTGCAGGACACAGATGCTGAGCATGTGCAACAGTTCACAC	1320
QY	2047	CTCCTGACAGCTCCATTTGATGTATCCAGAGCTGGGAGAGAAAGATTCGGGATTTACT	2106
Db	1321	CTTCTGACGCGCTCCATTTGACGTGTCCAGAGCTGGGAGAGAAAGATTCAGGATTTACT	1380
QY	2107	GATCTCCCAAGAGATCAGACATTTACTTATTTGATCAGCCTTTTGGAGCTGTTTGC	2166
Db	1381	GATCTCCCAAGAGATCAGACGTTTACTTATTTGATCAGCCTTTTGGAGCTGTTTGT	1440
QY	2167	CTCAGACTTTCCATCAGGTCACAACTGCTGAAGATAAGTTTGTCTTCTGCAATGACTT	2226
Db	1441	CTTAGACTTTCCATCAGGTCACAACTGCTGAAGATAAGTTTGTCTTCTGCAATGACTT	1500
QY	2227	GTCTGTCATCGACTTCAGTGCCTTCTGAGATTTGGGAGTGGCTCGACTCTATTAAAGAC	2286
Db	1501	GTCTGTCATCGACTTCAGTGCCTTCTGAGATTTGGGAGTGGCTCGACTCCATTAAAGAC	1560
QY	2287	TTTTCTTAAATTTGCAGAGCCTGAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAGCA	2346
Db	1561	TTTTCTTAAATTTGCAGAGCCTGAACCTTGATATCCAAAGCCTTAGCCTGCTGTCAGCA	1620
QY	2347	CTGAGCATGATCACAGAAAGACATGGGTTTAAAGAACCAAGAGAGTCGAAGAGCTATGC	2406
Db	1621	CTGAGTATGATCACAGAGCGACATGGTTTAAAGAACCAAGAGAGTCGAAGAGCTATGC	1680
QY	2407	AACAAGATCACAAGCAGTTTAAAGAACCAAGAGTAAGGACAGGCTCGAGGCCAC	2466
Db	1681	ACCAAGATCACAAGCAGTTTAAAGAACCAAGAGTAAGGACAGGCTCGAGGCCCTCG	1740
QY	2467	GAGTCAAGGCTCTGGTCCCTGTTAGACTGAGGAAGATCTGCACCTTGGGCTCCAG	2526
Db	1741	GAGCTCAAGGCTCTGGTCCCTGTTAGACTGAGGAAGATCTGCACCTTGGGCTCCAG	1800
QY	2527	CGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGACAAGCTC	2586
Db	1801	CGCATCTTCTACCTGAAGCTGAGGACTTGGTGTCTCCACCTTCCATCATTTGACAAGCTC	1860
QY	2587	TTCTTGGACACCTTCTTCTAA	2610
Db	1861	TTCTTGGACACCTTCTTCTGA	1884
RESULT 14	AF191211	1884 bp mRNA linear ROD 06-MAR-2000	
LOCUS	AF191211	Mus musculus orphan nuclear receptor TEC long isoform (Tec) mRNA, complete cds.	
DEFINITION	AF191211		
ACCESSION	AF191211		
VERSION	AF191211.1	GI:6176563	
KEYWORDS		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 1884)	
TITLE		Maltais, A. and Labelle, Y.	
JOURNAL		Structure and expression of the mouse gene encoding the orphan nuclear receptor TEC	
MEDLINE		DNA Cell Biol. 19 (2), 121-130 (2000)	
PUBMED		20163713	
REFERENCE		10701778	
AUTHORS		2 (bases 1 to 1884)	
TITLE		Maltais, A. and Labelle, Y.	
JOURNAL		Submitted (30-SEP-1999) Research Center, Pavillon St-Francois d'Assise, 10, rue de l'Esplanay, Quebec, Qc G1L 3L5, Canada	

FEATURES	Location/Qualifiers	Query Match	39.2%; Score 1485.6; DB 10; Length 1884;
source	1..1884	Best Local Similarity	87.2%; Pred. No. 1.1e-231;
gene	/organism="Mus musculus"	Matches 1642; Conservative	0; Mismatches 239; Indels 3; Gaps 1;
CDS	/mol_type="mRNA"		
	1..1884		
	/db_xref="taxon:10090"		
	/gene="Tec"		
	1..1884		
	/gene="Tec"		
	/note="similar to NOR-1, MINOR, CHN"		
	/codon_start=1		
	/product="orphan nuclear receptor TEC long isoform"		
	/protein_id="AAF05622.1"		
	/db_xref="GI:6176564"		
	/translation="MPCVQAQSPSPGPGSTVATQYGYSEYTTETINPNVYTKLMDLGS		
	THHATATTSLPSTFMEGYPSCSLYOMPPSPGPRELIKHREGRHGYHHH		
	HHHHHHHHQSQPSIPSPSGDEVLPSSTMYFKQSPSTPTTFCFPQAQALWDD		
	LPSPGCTAPGLDPOKAVPMAAARFPFFKPPSPHPPAPGAGHILGYDPTA		
	AAALSLPIGAAGAGSQAALLEGHPYGLPLAKRTATLTFFPLGLTASPTASSLGGSP		
	SLPSPNRSSEGGTCAVCGDNACQHYGVRTCEGKGFRTVQKNKXYYCLANKN		
	CPVDKRRNCOYCRFQKLSVGMKEVVRTDLSKRRGLRPSKPSPLQSPSPSP		
	SPPTCMNVALVRLTDAITPRDLDSYRPTDQATGDAEHWQDFNYLLTASIDVSR		
	SWAEKIPGFTDLPEKDOTLLIESAFLEFLVRLSIRSNTEADKFKCNGLVLRQL		
	RGFGEWLSIKDFSLNLQSLNDIOALACLALSMTIRHGLKEPKRVELCTKITS		
	LKDHRKQSALEPSEPKVLRALVELRKICTQGLQRIFYLKLEDLVPPSPVIDKLFD		
	LPF"		
ORIGIN			
		QY	730 ATGCGCTCGCTCCAGCCCAATATAGCCCTTCCCTTCCAGGTTCCAGTTATGCGGCGCAG 789
		Db	1 ATGCGCTCGCTCCAGCCCAATATAGCCCTTCCCTTCCAGGTTCCAGTTATGCGGCGCAG 60
		QY	790 ACATACAGCTCGGATATACACGAGAGATCATGACCCCGAGTACACCAAGCTGACCATG 849
		Db	61 ACTTATGCTCGGATATACACCAAGAAATCATGAACCCCGAGTACACCAAGCTGACCATG 120
		QY	850 GACCTTGGCAGCTAGATACGCGTACAGCACACAGCTCCCTGCCAGCAGCTAGTACC 909
		Db	121 GACCTCGGTAGACCGGGATCATGCCACCGCCACTACATCCCTGCCAGCTTCACTACC 180
		QY	910 TTCGTGGAGGGTACTCGAGCAACTACGAACTCAAGCCCTTCTCGGTGTACCAATGCGAG 969
		Db	181 TTCATGGAGGGTACCCCGAGAGCTGCGAACTCAAGCCCTTCTCGGTGTACCAATGCGAG 240
		QY	970 CGGCGCTTG--ATCAAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACCATCACCAC 1026
		Db	241 CTTCTGGGCTTGGGCTTGGATCAAGATGGAAGAGGGTCCGAGCATGGTACCCACAC 300
		QY	1027 CACCAACACACACACACACCATCACAGAGCAGCAGTACAGCAGCAGTCCATTCCTCCA 1086
		Db	301 CACCATCACCATCACCATCATCACACACACAGCAGCAGCAGCGCTTCCATTCCTCT 360
		QY	1087 GCCTCCAGCCCGAGGAGAGGTGTGCCCCAGCACTCCATGTACTTCAAGCAGTCCCCA 1146
		Db	361 CCCTCCGGCCCCGAGGAGGAGTGTGCCCCAGCACTCCATGTACTTCAAGCAGTCTCG 420
		QY	1147 CGGTCCACCCACACACCGCGCTTCCCGCGAGGGGGGGGGCTTATGGACGAGGCA 1206
		Db	421 CGGTCTACACCGACCACTCCAGGCTTCCCGCGAGGGGGGGGGCTGTGGGACGAGAG 480
		QY	1207 CTGCGCTCGGGCGCGGCTGTGATCGACCCCGCGCTGTGGACCCCGCGATGAAGGG 1266
		Db	481 CTGCGCTCTGGCGCTTGGTGTGATCGCTCGGAGACCGCTGTGGACCCCGAGATGAAGGG 540
		QY	1267 GTCCCCAGGTGGCGGCGCGCTTCCCGCTTCTTCCACTTCAAGCCCTTCCCGCGCAT 1326
		Db	541 GTACCCCGCATGGCGCTGCTGCGCGCTTCCCGCTTCTTCAAGCCCTTCCCGCGCAT 600
		QY	1327 CCCCCGGCGCCAGCCCGCGCGGCGCCACCTTGGGCTACGACCCGAGCGCGCTGCC 1386



Db 601 CCTCCGCGCCAGTCCAGCGCGGCCACCACTCGGCTATACCCCAACGCGCGAGCT 660  
Qy 1387 GGGCTCAGCTCGCGTGGAGCCGAGCGCGCGGGGAGCCAGCGCGCGCGCTTGA 1446  
Db 661 GCACTAGTCTGCGCTGGAGCGCGCGCGCGAGGAGGAGCAAGCTGCTCGCTCGAG 720  
Qy 1447 AGCCACCCGCTAGCGGCTGCGCTGGCAAGAGGGGGCGCCCGCTTCCCGCTCTC 1506  
Db 721 GGCCACCCATACGGGCTCGGCTGGCCAGAGAGCGGCCACGCTGACCTTCCCTCGCTG 780  
Qy 1507 GGCCTCAGCCCTCCCTTACCGGCTCAGCCTGTGGGCGAGAGTCCAGCCTGCGCTG 1566  
Db 781 GGCCTCAGAGCCTCCCGCAGCGCTGCTGGAGAGAGAGCGCCAGCGCTCCCATCG 840  
Qy 1567 CGCCCGCAGCAGGAGCTCGCTGCTGGCGAGGCGAGTGTGCGCTGTGGGGACACCGC 1626  
Db 841 CCACCCATAGAGCTCATCTCTGGGAGGCGCATGTGCGCTGTGGCGGACACAGCT 900  
Qy 1627 GCTGCCAGCACTACGGGCTGCGAAGCTGCGAGGGCTGCAAGGGCTTTTCAAGAGACA 1686  
Db 901 GCTGCCAGCACTACGGAGTCCGACCTGCGAGGGCTGCAAGGGCTTCTTCAAGAGAGC 960  
Qy 1687 GTGCAGAAAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCCGCTAGCAGAGA 1746  
Db 961 GTGCAGAAAAATGCAAAATATGTTTGCCTGGCAATATAAACTGCCCGCTAGCAGAGA 1020  
Qy 1747 CQTGCAAAACCGATGTCAGTACTGCTGATTTCAAGAGTGTCTCAGTGTGGAGTGTAA 1806  
Db 1021 CCGCGAAACCGATGTCAGTACTGCGATTTCAAGAGTGTCTCAGTGTGGAGTGTAA 1080  
Qy 1807 GAAGTTGTCCGTCAGATAGTCTGAAAGGGAGGAGGTGCTGCTGCCCTTCAAAACCAAG 1866  
Db 1081 GAAGTTGTCCGTCAGAGAGTCTGAAAGGGAGGAGGTGCTGCTGCCCTTCAAAACCAAG 1140  
Qy 1867 AGCCATTACACAGAGACCTTCTCAGCGCTCTCAGCTTCTCCTCCTCAATCTGATGATG 1926  
Db 1141 AGCCCATACACAGAGAGCCCTCGAGCGCTCTCCGCGCATCTCTCCGATCTGTATGATG 1200  
Qy 1927 AATGCGCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGTATTATCCAGATAC 1986  
Db 1201 AATGCGCTTGTCCGAGCTTTAAACAGATGCAACACCCAGAGATCTTGTATTATCCAGATAC 1260  
Qy 1987 TGTCCACTGACAGGCTGCTCAGGCGACAGATGCTGAGCATGTGCAACAAATTTTACAA 2046  
Db 1261 TGTCCCAACGACAGGCGCACTGAGGCGACAGATGCTGAGCATGTGCAACAGTTTCTACA 1320  
Qy 2047 CTCCTGACAGCTCCATTGATGATATCCAGAGCTGGCAGAGAAAGATTCGGGATTTACT 2106  
Db 1321 CTCCTGAGCGCTCCATTGAGCTGTCAGAGCTGGCAGAGAAAGATTCAGGATTTACT 1380  
Qy 2107 GATCTCCCCAAGAGATCAGACATTAATTATTAATCAGCTTTTGGAGCTGTTTCTC 2166  
Db 1381 GATCTCCCCAAGAGATCAGAGCTTACTATAGATCAGCTTTTGGAGCTGTTTCTC 1440  
Qy 2167 CTCAGACTTTCATCAGGTCAAACTGCTGAAGATAAGTTTGTCTTCTGCAATGCACTT 2226  
Db 1441 CTTAGACTTTCATCAGGTCAAACTGCTGAAGATAAGTTTGTCTTCTGCAATGCACTT 1500  
Qy 2227 GTCTGTCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTTATTAAGAC 2286  
Db 1501 GTCTGTCATCGACTTCAGTGCCTTCGAGATTTGGGAGTGGCTCGACTTATTAAGAC 1560  
Qy 2287 TTTTCTTTAAATTTGAGAGCTGCAACCTTGATATCCAGCTTACGCTGCTGTGAGCA 2346  
Db 1561 TTTTCTTTAAATTTGAGAGCTGCAACCTTGATATCCAGCTTACGCTGCTGTGAGCA 1620  
Qy 2347 CTGAGCATGATCAGAAAGACATCGGTTAAAGAACCAAGAGAGTCCAGAGAGCTATGC 2406  
Db 1621 CTGAGTATGATCAGAGAGCAGATCGGTAAAGAACCAAGAGAGTGGAGGAGCTATGC 1680  
Qy 2407 AACAGATCAGAGAGTAAAGAACCAAGAGTAAAGGAGAGGCTCTGAGGCCAC 2466

Db 1681 ACCAAGATCAACAAGCAGCTTAAAGACCAACAGAGAGGAGCAGGCTCTGGAGCCCTCG 1740  
Qy 2467 GAGTCCAAAGTCTCTGGTGGCTGTAGAACTGAGAAAGATCTGACCTGGGCTCCAG 2526  
Db 1741 GAGCTTAAAGTCTCTGGCGCTGTAGAACTGAGAAAGATCTGACCCAGGCGCTCCAG 1800  
Qy 2527 CGCATCTTCTTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCTTCAACAAGTCT 2586  
Db 1801 CGCATCTTCTTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCTTCAACAAGTCT 1860  
Qy 2587 TTCTTGGACACCTTACCTTTCTAA 2610  
Db 1861 TTCTTGGACACCTTCTTCTGA 1884

RESULT 15  
ALJ58937

LOCUS  
DEFINITION  
Human DNA sequence from clone RP11-6013 on chromosome 9, complete sequence.

ACCESSION  
ALJ58937VERSION  
ALJ58937.16 GI:16973850KEYWORDS  
HTG.SOURCE  
Homo sapiens (human)ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (Bases 1 to 146889)AUTHORS  
Pearce, A.TITLE  
Direct SubmissionJOURNAL  
Submitted (16-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Nov 17, 2001 this sequence version replaced gi.16304457.  
During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; Swi,  
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP  
database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping  
Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9

RP11-6013 is from the library RPCI-11.1 constructed by the group of  
Piet de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone  
RP11-6013 it may be shorter because we sequence overlapping  
sections only once, except for a short overlap.

The true right end of clone RP11-6013 is at 146889 in this  
sequence. The true left end of clone RP11-124M17 is at 93922 in  
this sequence. The true right end of clone RP11-30L7 is at 2000 in  
this sequence.

FEATURES  
source

Location/Qualifiers

1. 146889  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="9"

/clone="Rp11-6013"  
/clone\_lib="RPCI-11.1"

ORIGIN

Query Match 37.8%; Score 1435; DB 9; Length 146889;  
Best Local Similarity 100.0%; Pred. No. 1.8e-223;  
Matches 1435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2360 CAGAAAGACATGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCACAA 2419  
DB 25495 CAGAAGACATGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCACAA 2554  
QY 2420 GCAGTTTAAAGACCAACAGAGTGAAGGACAGGCTCTGAGGCCACCGAGTCCAAGGTCC 2479  
DB 25555 GCAGTTTAAAGACCAACAGAGTGAAGGACAGGCTCTGAGGCCACCGAGTCCAAGGTCC 25614  
QY 2480 TGGGTGCCCTGTGAGAACTGAGGAAGATCTGCACCCCTGGGCTCCAGCGCATCTTCTACC 2539  
DB 25615 TGGGTGCCCTGTGAGAACTGAGGAAGATCTGCACCCCTGGGCTCCAGCGCATCTTCTACC 25674  
QY 2540 TGAAGCTGAAGACTTGGTGTCTCCACCTTCCATCTTGAAGACTCTTCTGGACACCC 2599  
DB 25675 TGAAGCTGAAGACTTGGTGTCTCCACCTTCCATCTTGAAGACTCTTCTGGACACCC 25734  
QY 2600 TACCTTTCTAATCAGGACAGTGGAGCAGTGAAGTGCCTCTCTCTAGCAGCTCTGCTTGC 2659  
DB 25735 TACCTTTCTAATCAGGACAGTGGAGCAGTGAAGTGCCTCTCTCTAGCAGCTCTGCTTGC 25794  
QY 2660 TACGACGAAAGGATAGTGTGGAAACCTATCAITTCCTGCTCTCTTAAAGAGGAAA 2719  
DB 25795 TACGACGAAAGGATAGTGTGGAAACCTATCAITTCCTGCTCTCTTAAAGAGGAAA 25854  
QY 2720 GCAGTCTCTGTAAGAGGAAAGACTTTCTTTTCTGGCTCTCTTCTTCAACCTTAA 2779  
DB 25855 GCAGTCTCTGTAAGAGGAAAGACTTTCTTTTCTGGCTCTCTTCTTCAACCTTAA 25914  
QY 2780 AGCCAGAAAACCTGACAGATATTGCTGGGGTGTGTTTTATATTAGGCAATGGGGA 2839  
DB 25915 AGCCAGAAAACCTGACAGATATTGCTGGGGTGTGTTTTATATTAGGCAATGGGGA 25974  
QY 2840 TGGGTGGGAGGGGTATAGTTCAATGAGGGTTTTCTAAGAAATGCTAACAAAGCATT 2899  
DB 25975 TGGGTGGGAGGGGTATAGTTCAATGAGGGTTTTCTAAGAAATGCTAACAAAGCATT 26034  
QY 2900 TTGGACAATGCTATCCAGCAGGAAAAAAGGATAATATACTGTTTTAAACTCTTTC 2959  
DB 26035 TTGGACAATGCTATCCAGCAGGAAAAAAGGATAATATACTGTTTTAAACTCTTTC 26094  
QY 2960 TGGGGAATCCAATTATAGTTGCTTTGTAATTAAGAACAGAACAGCAGGAGGTGTTCCG 3019  
DB 26095 TGGGGAATCCAATTATAGTTGCTTTGTAATTAAGAACAGAACAGCAGGAGGTGTTCCG 26154  
QY 3020 CAGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTATCAAGGTA 3079  
DB 26155 CAGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAAAATATGCTTCTGTATCAAGGTA 26214  
QY 3080 CGTATGTGTGCAACAAGGAGAACTTCTCTTTTAAATTTCTCTCTCTCTTTTAACT 3139  
DB 26215 CGTATGTGTGCAACAAGGAGAACTTCTCTTTTAAATTTCTCTCTCTCTTTTAACT 26274  
QY 3140 AATAGTGAAGATGGAGGATTTACCTACAAATCAGACATGGCAACAAATATGGCTGT 3199  
DB 26275 AATAGTGAAGATGGAGGATTTACCTACAAATCAGACATGGCAACAAATATGGCTGT 26334  
QY 3200 TGCTTCCATAAACAAGTGAATTTTTTAAAGTGTGCTTACTAAGTCTTGTATTAACT 3259  
DB 26335 TGCTTCCATAAACAAGTGAATTTTTTAAAGTGTGCTTACTAAGTCTTGTATTAACT 26394  
QY 3260 TCTCTTTTATCTATATGAATAAAGAGGAGTCAATGTTAGCAATGACAGTTAA 3319  
DB 26395 TCTCTTTTATCTATATGAATAAAGAGGAGTCAATGTTAGCAATGACAGTTAA 26454  
QY 3320 TATCCCTAGCAGGCTGTGTTCACTTCCCTGTGATCCCTTCTGAGGTATGGCCCATC 3379

DB 26455 TATCCCTAGCAGGCTGTGTTCACTTCCCTGTGATCCCTTCTGAGGTATGGCCCATC 26514  
QY 3380 CAAGACTTTTAGGCCATTCTTGATGGAACACAGATCCCTGCCCTGACTGTCCAGCTATCCT 3439  
DB 26515 CAAGACTTTTAGGCCATTCTTGATGGAACACAGATCCCTGCCCTGACTGTCCAGCTATCCT 26574  
QY 3440 GAAAGTGGATCAGATTATAAACTGGATTACATGTAAGTGTGTTGTTGTTCTATCAAC 3499  
DB 26575 GAAAGTGGATCAGATTATAAACTGGATTACATGTAAGTGTGTTGTTGTTCTATCAAC 26634  
QY 3500 CCCACACAGATTCCCTAAACTTGTCTTCACTTATAGTAAGTGTGTTTATATATATATCA 3559  
DB 26635 CCCACACAGATTCCCTAAACTTGTCTTCACTTATAGTAAGTGTGTTTATATATATCA 26694  
QY 3560 AGCGCCATAAGTCACTTGAATATTTGATCCCTTAGATAAGAAATGCAAAATCAGCAGGAAC 3619  
DB 26695 AGCGCCATAAGTCACTTGAATATTTGATCCCTTAGATAAGAAATGCAAAATCAGCAGGAAC 26754  
QY 3620 TGGTCATACAGGTTAAGCACCAGGACCAATAGGATTTTATAGATATATATATATAT 3679  
DB 26755 TGGTCATACAGGTTAAGCACCAGGACCAATAGGATTTTATAGATATATATATATAT 26814  
QY 3680 GTTATTGGTTAAGGAGACAAATTTTGGAGAGCAAGCAATCTTTTAAAAAATAGTATGAA 3739  
DB 26815 GTTATTGGTTAAGGAGACAAATTTTGGAGAGCAAGCAATCTTTTAAAAAATAGTATGAA 26874  
QY 3740 TGTGAATACTAGAAAAAGATTTAAAAAATAGTATGATGAGTGTGAGTACTAGGAAGGAT 3794  
DB 26875 TGTGAATACTAGAAAAAGATTTAAAAAATAGTATGATGAGTGTGAGTACTAGGAAGGAT 26929

Search completed: March 21, 2005, 16:56:22  
Job time : 15982 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 21, 2005, 04:05:50 ; Search time 1901 Seconds  
(without alignments)  
11814.571 Million cell updates/sec

Title: US-10-608-863-1  
Perfect score: 3794  
Sequence: 1 ataatgacgtccgagaga.....gtgtgagtactaggagagat 3794

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_16Dec04:\*
- 1: Geneseqn19808:\*
  - 2: Geneseqn19808:\*
  - 3: Geneseqn2000s:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2002bs:\*
  - 8: Geneseqn2003as:\*
  - 9: Geneseqn2003bs:\*
  - 10: Geneseqn2003cs:\*
  - 11: Geneseqn2003ds:\*
  - 12: Geneseqn2004as:\*
  - 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3794	100.0	3794	AD113007	Ad113007 Human NOR
2	3722	98.1	3802	ABK51689	Abk51689 Human nuc
3	3722	98.1	3802	ABK83698	Abk83698 Human cdn
4	3722	98.1	3802	AD157170	Ad157170 Human NOV
5	3679.2	97.0	5828	AD157170	Ad157170 Human cdn
6	3668.2	96.7	5642	ABX62937	Abx62937 Human act
7	3653.6	96.3	3762	AAT73334	Aat73334 Human neu
8	3352	88.4	4162	ADP28866	Adp28866 Human nor
9	2875	75.8	4977	ABL61892	Ab161892 Colon ade
10	2875	75.8	4977	ABK84229	Abk84229 Human cdn
11	2875	75.8	4977	ADL83261	Adl83261 Human PRO
12	2875	75.8	4977	ADL83261	Adl83261 Human NF-
13	2689.6	70.9	2714	AD131602	Ad131602 Human cdn
14	2500	65.9	4400	AAT16151	Aat16151 Apoptotic
15	2500	65.9	4400	ABN85437	Abn85437 Murine Ne
16	2500	65.9	4400	ADP28867	Adp28867 Rat nor-1
17	2371	62.5	5115	ADP28873	Adp28873 Pig nor-1
18	2055	54.2	2055	ADP05710	Adp05710 Human nuc
19	1485.6	39.2	1884	ABN85436	Abn85436 Murine Ne
20	1022.8	27.0	1606	ADP28871	Adp28871 Pig nor-1

21	937.4	24.7	1290	12	ADP05712	Adp05712 Mouse nuc
22	499.4	13.2	3471	12	ADQ22833	Adq22833 Human eof
23	497	13.1	2799	9	ADB84081	Adb84081 Mutant NU
24	497	13.1	3421	9	ADB84043	Adb84043 Human NUR
25	497	13.1	3427	2	AAQ56733	Aaq56733 CDNA sequ
26	497	13.1	3427	4	AAS15573	Aas15573 DNA encod
27	497	13.1	3427	9	ADB84033	Adb84033 Human NUR
28	497	13.1	3427	10	ADH28985	Adh28985 Human chr
29	497	13.1	3427	10	ACC58210	Acc58210 Human NUR
30	497	13.1	3427	11	AD132074	Ad132074 Human cdn
31	497	13.1	3427	12	ADH75169	Adh75169 Human TIN
32	497	13.1	3427	13	ADP23739	Adp23739 PRO polyP
33	497	13.1	3436	10	ADE25697	Ad25697 Human cdn
34	496	13.1	2154	9	ADB84053	Adb84053 Rat NURR1
35	495.4	13.1	1797	4	AAC67273	Aac67273 Human nuc
36	495.4	13.1	1797	12	ADP05706	Adp05706 Human nuc
37	492.8	13.0	2002	13	ADR14978	Adr14978 Rat elect
38	492.4	13.0	2247	9	ADB84058	Adb84058 Murine NU
39	492.4	13.0	2458	9	ADB84032	Adb84032 Murine NU
40	490.8	12.9	1797	12	ADP05708	Adp05708 Mouse nuc
41	486.4	12.8	1821	5	AAS14641	Aas14641 Rat Nurr1
42	486	12.8	1821	5	AAS14641	Aas14641 Mouse cdn
43	483.2	12.7	2559	10	ACC58209	Acc58209 Rat Nurr1
44	469.4	12.4	2469	10	ADF28894	Adf28894 Human nur
45	469.4	12.4	2469	10	ADK61137	Adk61137 Ovarian c

ALIGNMENTS

RESULT 1  
AD113007  
ID AD113007 standard; cdna; 3794 BP.  
XX  
AC AD113007;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Human NOR-1 (MINOR) cdna sequence SeqID 1.  
XX  
KW human; gene; ss; allergic disease; NOR-1; MINOR; eosinophil;  
KW atopic dermatitis; antiallergic; antiinflammatory; dermatological.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 730..2610  
FT /\*tag= a  
FT /product= "NOR-1 protein"  
XX  
PN WO2004003198-A1.  
XX  
PD 08-JAN-2004.  
XX  
PF 27-JUN-2003; 2003WO-JP008199.  
XX  
PR 27-JUN-2002; 2002JP-00188490.  
XX  
PA (GENO-) GENOX RES INC.  
PA (NIGE-) JAPAN GEN AGENCY NATION.  
XX  
PI Hashida R, Kagaya S, Yayoi Y, Sugita Y, Saito H;  
XX  
DR WPI; 2004-083057/08.  
DR P-PSDB; AD113008.  
XX  
PT Examining allergic diseases e.g. atopic dermatitis by differential  
PT display based on gene expression of NOR-1 receptor protein, also  
XX applicable in screening compounds for treatment of allergic diseases.  
XX  
PS Example 1; SEQ ID NO 1; 155pp; Japanese.  
XX  
CC This invention relates to a novel method for examining allergic diseases

that comprises comparing the expression levels of a gene encoding the NOR-1 receptor protein between patients and healthy individuals. Specifically, the NOR-1 gene, also referred to as MNOR, is expressed in the specialist white blood cells known as eosinophils and is involved in mediating an allergic reaction. The present invention describes a differential display method that can identify the expression level of this gene in order to identify its usefulness in diagnosing allergic diseases such as atopic dermatitis. Furthermore, compositions can also be used to screen compounds for the treatment of allergic diseases. Accordingly, they exhibit various activities including antiallergic, antiinflammatory and dermatological. This polynucleotide sequence is the human NOR-1 cDNA sequence of the invention.

SQ	Sequence	3794 BP; 917 A; 1140 C; 910 G; 827 T; 0 U; 0 Other;
	Query Match	100.0%; Score 3794; DB 12; Length 3794;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 3794; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATAAATGACGTCCGAGAGCGAGCGAAGCGCAGCGCGGAGCGGAGCGGAGTCTCTCGCCT 60
Db	1	ATAAATGACGTCCGAGAGCGAGCGAAGCGCAGCGCGGAGCGGAGCGGAGTCTCTCGCCT 60
Qy	61	CCGCGCCCCACCCCTCCAGTCTCTGCTCTCTCTCCGCTCCGCTCCCATATACAGAGCGCGTCA 120
Db	61	CCGCGCCCCACCCCTCCAGTCTCTGCTCTCTCTCCGCTCCCATATACAGAGCGCGTCA 120
Qy	121	CACCGGCTCCCTCACTCGGCACACAGACACAAGCGCGGCACACAGGCTCCGCACACACAC 180
Db	121	CACCGGCTCCCTCACTCGGCACACAGACACAAGCGCGGCACACAGGCTCCGCACACACAC 180
Qy	181	TTTCGCTCTCCCGGGCGCTACACCCCTCTTCGCTGAGCCCTTTCGCGGTGACGGCGGG 240
Db	181	TTTCGCTCTCCCGGGCGCTACACCCCTCTTCGCTGAGCCCTTTCGCGGTGACGGCGGG 240
Qy	241	CCGAGCTGGAGCGCCCTCCCGGGCTCACTTTTGCAACGCTGACGGTCCCGGCAGTGGCGCG 300
Db	241	CCGAGCTGGAGCGCCCTCCCGGGCTCACTTTTGCAACGCTGACGGTCCCGGCAGTGGCGCG 300
Qy	301	TGGAGTGGGAACAGCGGGGGGATCTCTCCCGCTTGTGTCAAGCCCAAGCCAGAGACGCCCG 360
Db	301	TGGAGTGGGAACAGCGGGGGGATCTCTCCCGCTTGTGTCAAGCCCAAGCCAGAGACGCCCG 360
Qy	361	CGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCGCAGCGCGCT 420
Db	361	CGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCGCAGCGCGCT 420
Qy	421	CACCGCTCCGGGAGCCGCTGGGCTTGTACACCGCAGCCCTTCGGGACAGCAGCTGTGA 480
Db	421	CACCGCTCCGGGAGCCGCTGGGCTTGTACACCGCAGCCCTTCGGGACAGCAGCTGTGA 480
Qy	481	CTCCCCCGCAGTGCAGATTTCCGAGACGCTCTCTAGAAACTCGCTCTTAAAGACGGAACCG 540
Db	481	CTCCCCCGCAGTGCAGATTTCCGAGACGCTCTCTAGAAACTCGCTCTTAAAGACGGAACCG 540
Qy	541	CCACGACACTCAAAAGCCCACTTCGGAAGAGGGCAGCCGCGAAGCCCGGCCCTTGAGCCT 600
Db	541	CCACGACACTCAAAAGCCCACTTCGGAAGAGGGCAGCCGCGAAGCCCGGCCCTTGAGCCT 600
Qy	601	GGACCTTTAGCGGTCCGGCGACGACTCCCGCGCTTTGCGCTCGCCGGAGCTCCCGCTCCT 660
Db	601	GGACCTTTAGCGGTCCGGCGACGACTCCCGCGCTTTGCGCTCGCCGGAGCTCCCGCTCCT 660
Qy	661	CCTACACTCTCAGCCTCGCTGGAGAGACCCCGACCCACCATTTCAGCGCGCAAGATAC 720
Db	661	CCTACACTCTCAGCCTCGCTGGAGAGACCCCGACCCACCATTTCAGCGCGCAAGATAC 720
Qy	721	CTTCCAGATATGCGCTCGCTCAAGCCCAATATAGCCCTTCCCGCTCCAGGTTCCAGTTAT 780
Db	721	CTTCCAGATATGCGCTCGCTCAAGCCCAATATAGCCCTTCCCGCTCCAGGTTCCAGTTAT 780
Qy	781	GGGGCGCAGACATACAGTTCGGAATACACCGAGAGATCATGAACCCCGACTACACCAAG 840

1921 ATGATGAATGCGCTTGTCCGAGCTTTAAACAGACTCAACACCAGAGATCTTGATTTATCC 1980  
1921 ATGATGAATGCGCTTGTCCGAGCTTTAAACAGACTCAACACCAGAGATCTTGATTTATCC 1980  
1981 AGATACTGTCCCACTGACCAAGGCTGTGCGAGGACACAGATGCTGAGCATGTGCAACAATTC 2040  
1981 AGATACTGTCCCACTGACCAAGGCTGTGCGAGGACACAGATGCTGAGCATGTGCAACAATTC 2040  
2041 TACAACCTCTGACAGCTCCATTTGATGTATCCAGAACTGCGGAGAGATTCGGGA 2100  
2041 TACAACCTCTGACAGCTCCATTTGATGTATCCAGAACTGCGGAGAGATTCGGGA 2100  
2101 TTTACTGATCTCCCAAGAGATCAGATTAATTTGATTAATTTGAGATCAGCCCTTTTGGAGCTG 2160  
2101 TTTACTGATCTCCCAAGAGATCAGATTAATTTGATTAATTTGAGATCAGCCCTTTTGGAGCTG 2160  
2161 TTTGCTCTCAGACTTTCCATCAGGTCAAACTGCTGAAGATAAAGTTTGTGTTCTGCAAT 2220  
2161 TTTGCTCTCAGACTTTCCATCAGGTCAAACTGCTGAAGATAAAGTTTGTGTTCTGCAAT 2220  
2221 GGACTTGTCTGATCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATT 2280  
2221 GGACTTGTCTGATCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATT 2280  
2281 AAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTGATATCCAGCCCTTAGCCTGCTG 2340  
2281 AAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTGATATCCAGCCCTTAGCCTGCTG 2340  
2341 TCAGACTGAGCATGATCAGAGAAAGATGCTGTTTAAAGAACCCAAAGAGAGTTCGAAGAG 2400  
2341 TCAGACTGAGCATGATCAGAGAAAGATGCTGTTTAAAGAACCCAAAGAGAGTTCGAAGAG 2400  
2401 CTATGCAACAAGATCAACAGAGTTTAAAGAACCCAGAGTAAGGGACAGGCTCTGGAG 2460  
2401 CTATGCAACAAGATCAACAGAGTTTAAAGAACCCAGAGTAAGGGACAGGCTCTGGAG 2460  
2461 CCCACCGAGTCCAAGGCTCGGGTGCCTGGTAGAAGTCTGAGAGTCTGACCCCTGGGC 2520  
2461 CCCACCGAGTCCAAGGCTCGGGTGCCTGGTAGAAGTCTGAGAGTCTGACCCCTGGGC 2520  
2521 CTCACGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATTGAC 2580  
2521 CTCACGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATTGAC 2580  
2581 AAGCTCTTCTGACACACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGGCTGCTCC 2640  
2581 AAGCTCTTCTGACACACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGGCTGCTCC 2640  
2641 TCTCTAGCACCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTTATCATTTCTCTG 2700  
2641 TCTCTAGCACCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTTATCATTTCTCTG 2700  
2701 TCTTCTCTTAAGAGGAAAGCAGCTCTCTGAGAAAGCAAGACTTCTTTTTTTCTGGC 2760  
2701 TCTTCTCTTAAGAGGAAAGCAGCTCTCTGAGAAAGCAAGACTTCTTTTTTTCTGGC 2760  
2761 TCTTTTCTTACAACTTAAGCCAGAGAACTTGCAGAGTATTGTGTGGGGTGTGTTTT 2820  
2761 TCTTTTCTTACAACTTAAGCCAGAGAACTTGCAGAGTATTGTGTGGGGTGTGTTTT 2820  
2821 ATATTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTTAAGA 2880  
2821 ATATTAGGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTTAAGA 2880  
2881 AATTGCTACAAAGCACTTTTGGACAATGCTATCCAGCAGGAAAGAAAGGATATATA 2940  
2881 AATTGCTACAAAGCACTTTTGGACAATGCTATCCAGCAGGAAAGAAAGGATATATA 2940  
2941 ACTGTTTAAAACTCTTCTGGGGAATCCAAATATAGTTGCTTTGTATTAAAAACAAGA 3000  
2941 ACTGTTTAAAACTCTTCTGGGGAATCCAAATATAGTTGCTTTGTATTAAAAACAAGA 3000

3001 ACAGCCAAAGGGTTGTTCCGCCAGGCTAGGATGTCTTAAAGATTTGGTCCCTTGAAATAT 3060  
3001 ACAGCCAAAGGGTTGTTCCGCCAGGCTAGGATGTCTTAAAGATTTGGTCCCTTGAAATAT 3060  
3061 GCTTCTCTGATCAAAAGGTACGTATGTGTGTCAAAACAAAGCAGAACTCTCTTTAAATTC 3120  
3061 GCTTCTCTGATCAAAAGGTACGTATGTGTGTCAAAACAAAGCAGAACTCTCTTTAAATTC 3120  
3121 CTTCTCTCTTTATTTTAAACAAATGGTGAAGATGGAAGATTACCTACAAATCAGACATGG 3180  
3121 CTTCTCTCTTTATTTTAAACAAATGGTGAAGATGGAAGATTACCTACAAATCAGACATGG 3180  
3181 CAAAACAATAAATGGCTTTGCTTCCATAAACAAGTGCATTTTAAAAGTGTGCTCTTA 3240  
3181 CAAAACAATAAATGGCTTTGCTTCCATAAACAAGTGCATTTTAAAAGTGTGCTCTTA 3240  
3241 CTAAGTCTTGTGTTTAACTCTCCCTTTATTTCTATATGGAATAAAGAGGAGGAGTCATG 3300  
3241 CTAAGTCTTGTGTTTAACTCTCCCTTTATTTCTATATGGAATAAAGAGGAGGAGTCATG 3300  
3301 TTAGCAAAATGACACAGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTGCAATCC 3360  
3301 TTAGCAAAATGACACAGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTGCAATCC 3360  
3361 TTCTGAGGTATGGCCCATCCCAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCC 3420  
3361 TTCTGAGGTATGGCCCATCCCAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCC 3420  
3421 CTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGT 3480  
3421 CTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGT 3480  
3481 TTGTTGTTGTTCTATCAACCCCAAGAGTTCCTTAACTTGTGTTTGGAGAGCAAGCAATCT 3540  
3481 TTGTTGTTGTTCTATCAACCCCAAGAGTTCCTTAACTTGTGTTTGGAGAGCAAGCAATCT 3540  
3541 ACTGCTATATTTCATTCAGAGCGCCATTAAGTCAAGTGTGATTTGATCCCTAGATAAGAA 3600  
3541 ACTGCTATATTTCATTCAGAGCGCCATTAAGTCAAGTGTGATTTGATCCCTAGATAAGAA 3600  
3601 CATGCAAAATCAGCAGGAACTGGTCTATACAGGTTAAGCAGCAGGACAAATAGGATTTTAA 3660  
3601 CATGCAAAATCAGCAGGAACTGGTCTATACAGGTTAAGCAGCAGGACAAATAGGATTTTAA 3660  
3661 TAGATATAATTTAAATTTTGTGTTTAAAGAGCAAAATTTTGGAGAGCAAGCAATCT 3720  
3661 TAGATATAATTTAAATTTTGTGTTTAAAGAGCAAAATTTTGGAGAGCAAGCAATCT 3720  
3721 TTTTAAAAATAGTATGAATGTAATCTAGAAAAAGATTTAAAAATAGTATGAGTGTGA 3780  
3721 TTTTAAAAATAGTATGAATGTAATCTAGAAAAAGATTTAAAAATAGTATGAGTGTGA 3780  
3781 GTACTAGGAAGGAT 3794  
3781 GTACTAGGAAGGAT 3794

## RESULT 2

ABK51689

ID ABK51689 standard; DNA; 3802 BP.

XX ABK51689;

AC ABK51689;

XX 30-JUL-2002 (first entry)

XX Human nuclear receptor NOR1 gene sequence.

XX Human; nuclear receptor; NURR; inflammatory immune disease; arthritis;

XX corticotropin releasing hormone; receptor; CRH; rheumatoid arthritis;

XX chronic inflammatory joint disease; psoriatic arthritis; thyroiditis;

XX sarcoid arthritis; ulcerative colitis; NOR1; gene; ds.

XX Homo sapiens.

XX WO200187923-A1.  
PN 22-NOV-2001.  
PD 11-MAY-2001; 2001WO-US015311.  
XX 12-MAY-2000; 2000US-0203645P.  
XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX Murphy E, Conneely OM, Fitzgerald O, Bresihan B;  
PI WPI; 2002-075311/10.  
XX  
XX Treating inflammatory immune disease such as arthritis, comprises  
PT suppressing expression level of NURR subfamily of nuclear transcription  
PT factors, or corticotropin releasing hormone receptor.  
XX  
XX Claim 10; Page; 123pp; English.  
XX  
XX The present invention relates to a new method of treating an organism for  
CC an inflammatory immune disease. The method of the invention comprises  
CC reducing expression of a NURR subfamily nucleic acid sequence or  
CC corticotropin releasing hormone (CRH) receptor nucleic acid sequence,  
CC inhibiting transcriptional activity of a NURR superfamily member/CRH  
CC receptor amino acid sequence, or reducing the level of NURR superfamily  
CC member/CRH receptor sequence. The method is useful for treating an  
CC organism for an inflammatory immune diseases such as chronic inflammatory  
CC joint disease, preferably arthritis, selected from rheumatoid arthritis,  
CC psoriatic arthritis and sarcoid arthritis, ulcerative colitis and  
CC thyroiditis. The method is also useful for screening a compound that  
CC interferes with interaction of a NURR subfamily polypeptide with a  
CC ligand, or identifying a compound for the treatment of an inflammatory  
CC immune response. The agonist of the invention is useful for inhibiting  
CC transcriptional activity of nuclear receptor polypeptide and the  
CC antagonist is useful for decreasing the expression of a NURR subfamily  
CC member. The present nucleic acid sequence represents the human nuclear  
CC receptor NOR1 gene of the invention. Note: The present sequence is not  
CC shown in the specification but was obtained from Genbank (1651190)  
XX  
XX Sequence 3802 BP; 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;  
SQ  
Query Match 98.1%; Score 3722; DB 6; Length 3802;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;  
QY 1 ATAAATGAGCTCCGAGAGAGCGGACGCGACGCGGAGCGGAGTCTCCTGCT 60  
DB 1 ATAAATGAGCTCCGAGAGAGCGGACGCGACGCGGAGCGGAGTCTCCTGCT 60  
QY 61 CCGGCCCCCACCCTCCAGCTCCTGCTCCTCCTCCTCCGCTCCCATACAGACGGCTCA 120  
DB 61 CCGGCCCCCACCCTCCAGCTCCTGCTCCTCCTCCTCCGCTCCCATACAGACGGCTCA 120  
QY 121 CACCGCTCCCTCACTCGCACACACAGACACAAGCGCGCACAGCTCCG--CACACAC 178  
DB 121 CACCGCTCCCTCACTCGAACACACAGACACAAGCGCGCACAGGCTCCGACACACAC 180  
QY 179 ACTTCGCTCCCGCGGCTCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 238  
DB 181 ACTTCGCTCCCGCGGCTCACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 239 CCGCGAGCTGGAACGCCCCCTCCCGGCTCACTTTGCAACGCTGACGGTCCCGGCAAGTGGC 298  
DB 241 CCGCGAGCTGGAACGCCCCCTCCCGGCTCACTTTGCAACGCTGACGGTCCCGGCAAGTGGC 300  
QY 299 CGTGAGGTGGAAACAGCGGCGGCTATCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 358  
DB 301 CGTGAGGTGGAAACAGCGGCGGCTATCTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 359 CCGGGAACCTCTCGGCTGCTCTCCCATGAGTCGGGATCGGAGATCCCCACAGCGCG 418

DB 361 CGCGGAACCTCTCGGCTGCTCTCTCCCATGAGTCGGGATCGGAGATCCCCACAGCGG 420  
QY 419 CTCACCGCTCCGGGAGCGCTGGCTTGTACACCGCAGCGCTTCCGGGAGCAGCAGCTGT 478  
DB 421 CTCACCGCTCCGGGAGCGCTGGCTTGTACACCGCAGCGCTTCCGGGAGCAGCAGCTGT 480  
QY 479 GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAAAACCTGCTCTTAAGACGGAAC 538  
DB 481 GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAAAACCTGCTCTTAAGACGGAAC 540  
QY 539 CGGCACAGCACTCAAAAGCCCACTCGGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 598  
DB 541 CGGCACAGCACTCAAAAGCCCACTCGGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC 600  
QY 599 CTGACCCCTTAGCGGTGCGGGGAGCACTGCGGGGCTTTCGCTCGCGGAGCTCGGCTC 658  
DB 601 CTGACCCCTTAGCGGTGCGGGGAGCACTGCGGGGCTTTCGCTCGCGGAGCTCGGCTC 660  
QY 659 CTCTTACACTCTCAGCTCCGCTGAGAGAGCCCGCAGCCCACTTTCAGCGGCGAAGAT 718  
DB 661 CTCTTACACTCTCAGCTCCGCTGAGAGAGCCCGCAGCCCACTTTCAGCGGCGAAGAT 720  
QY 719 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 778  
DB 721 ACCCTCCAGATATGCCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 780  
QY 779 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACA 838  
DB 781 ATGCGGCGCAGACATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACA 840  
QY 839 AGCTGACCATGACCTTGGGAGCACTGAGATCAAGGCTACAGCCACACAGCTCCCTGCCA 898  
DB 841 AGCTGACCATGACCTTGGGAGCACTGAGATCAAGGCTACAGCCACACAGCTCCCTGCCA 900  
QY 899 GCATCAGTACCTTCTGAGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCTCGGCT 958  
DB 901 GCATCAGTACCTTCTGAGAGGCTACTCGAGCACTACGAACTCAAGCTTCTCTCGGCT 960  
QY 959 ACCAAATCAGCGGCTTGTATCAAGTGGAGGAGGCGGCGGCGCCAGCTACCATACC 1018  
DB 961 ACCAAATCAGCGGCTTGTATCAAGTGGAGGAGGCGGCGGCGGCGGCGGCTACCATACC 1020  
QY 1019 ATCACCAC 1078  
DB 1021 ATCACCAC 1080  
QY 1079 TTCTCTCAGCTCCAGCCCGGAGGAGGAGTGTGCTGCCAGCACCTTCCATGTTCAAGC 1138  
DB 1081 TTCTCTCAGCTCCAGCCCGGAGGAGGAGTGTGCTGCCAGCACCTTCCATGTTCAAGC 1140  
QY 1139 AGTCCCAACCGTCCACACCGGCTTCCCGGCGGCTTCCCGGCGGCGGCGGCTTATGGG 1198  
DB 1141 AGTCCCAACCGTCCACACCGGCTTCCCGGCGGCTTCCCGGCGGCGGCGGCTTATGGG 1200  
QY 1199 ACGAGGCACTGCCCTCCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1258  
DB 1201 ACGAGGCACTGCCCTCCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
QY 1259 TGAAGCGGCTCCCAACCGTGGCGGCGGCTTCCCGCTCTTCCATTTCAAGCCCTCGC 1318  
DB 1261 TGAAGCGGCTCCCAACCGTGGCGGCGGCTTCCCGCTCTTCCATTTCAAGCCCTCGC 1320  
QY 1319 CGCGCATCTCCCGCGGCTGAGCTCGGCGGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1378  
DB 1321 CGCGCATCTCCCGCGGCTGAGCTCGGCGGCGGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1380  
QY 1379 CGCTGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1438  
DB 1381 CGCTGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1440  
QY 1439 CGCTTGAAGAGCCACCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTTCC 1498  
DB 1441 CGCTTGAAGAGCCACCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTTCC 1500



1499 CGCCTCTCGGCTCAGGCCCTCCCTACCGGCTCAGGCTGCTGGCGAGAGTCCAGCC 1558  
1501 CGCCTCTCGGCTCAGGCCCTCCCTACCGGCTCAGGCTGCTGGCGAGAGTCCAGCC 1560  
1559 TGCCTGCGCGCCGAGCAGAGTCTGCTGCTGCGGAGGCGACGTGTCGCTGCGGG 1618  
1561 TGCCTGCGCGCCGAGCAGAGTCTGCTGCTGCGGAGGCGACGTGTCGCTGCGGG 1620  
1619 ACAAGCGCGCTGCCAGCACTACGCGTGCAGAACCTGCGAGGGCTGCAAGGGCTTTTCA 1678  
1621 ACAAGCGCGCTGCCAGCACTACGCGTGCAGAACCTGCGAGGGCTGCAAGGGCTTTTCA 1680  
1679 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGGCTGCGCAATAAAACTGCCAGTAG 1738  
1681 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGGCTGCGCAATAAAACTGCCAGTAG 1740  
1739 ACAAGAGAGCTGCAAAACCGATGTCAGTACTGTCGATTTTCAGAAAGTGTCTCAGTGTGGAA 1798  
1741 ACAAGAGAGCTGCAAAACCGATGTCAGTACTGTCGATTTTCAGAAAGTGTCTCAGTGTGGAA 1800  
1799 TGGTAAAAAGAGTTGTTCGATCAGATAGTCTGAAAGGGAGGAGAGTCTGTCGCCCTCCA 1858  
1801 TGGTAAAAAGAGTTGTTCGATCAGATAGTCTGAAAGGGAGGAGAGTCTGTCGCCCTCCA 1860  
1859 AACCAAGAGGCCATTTACAAAGGAACCTTCTCAGGCCCTCTCCACTTCTCCTCCAACT 1918  
1861 AACCAAGAGGCCATTTACAAAGGAACCTTCTCAGGCCCTCTCCACTTCTCCTCCAACT 1920  
1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGCTCAACACCCAGAGATCTTGATTA 1978  
1921 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGCTCAACACCCAGAGATCTTGATTA 1980  
1979 CCAGATACTGTCCCACTGACAGGCTGTGTCAGGCAAGATGTCGAGCATGTGCAACAAT 2038  
1981 CCAGATACTGTCCCACTGACAGGCTGTGTCAGGCAAGATGTCGAGCATGTGCAACAAT 2040  
2039 TCTACAACTCTGACAGCTCCATGATGATTCAGAAAGCTGGGAGGAGAAAGATTCGG 2098  
2041 TCTACAACTCTGACAGCTCCATGATGATTCAGAAAGCTGGGAGGAGAAAGATTCGG 2100  
2099 GATTTACTGATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCCTTTTGGAGC 2158  
2101 GATTTACTGATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCCTTTTGGAGC 2160  
2159 TGTGTGCTCTCAGACTTTCCATCAGGTCAAACACTGCTGAAGATAGTTTGTCTGCA 2218  
2161 TGTGTGCTCTCAGACTTTCCATCAGGTCAAACACTGCTGAAGATAGTTTGTCTGCA 2220  
2219 ATGGACTGTGCTGATCAGCTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2278  
2221 ATGGACTGTGCTGATCAGCTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2280  
2279 TTAAGAGCTTTTCCCTTAATTTGACAGGCTGAACTGATATCAAGCTTAGCCTGCC 2338  
2281 TTAAGAGCTTTTCCCTTAATTTGACAGGCTGAACTGATATCAAGCTTAGCCTGCC 2340  
2339 TGTGAGCACTGAGCATGATCAGAGAAAGACATGGGTTAAAGAAACCAAGAGAGTCTGAAG 2398  
2341 TGTGAGCACTGAGCATGATCAGAGAAAGACATGGGTTAAAGAAACCAAGAGAGTCTGAAG 2400  
2399 AGCTATGCAACAGATTCAGAGCAGTTTAAAGACCAAGAGTAAAGGAGAGGCTCTGG 2458  
2401 AGCTATGCAACAGATTCAGAGCAGTTTAAAGACCAAGAGTAAAGGAGAGGCTCTGG 2460  
2459 AGCCCAAGAGTCCAGGCTCTGGTGCCTGGTAGAAGTGGAGAGTCTGACCCCTGG 2518  
2461 AACCAAGAGTCCAGGCTCTGGTGCCTGGTAGAAGTGGAGAGTCTGACCCCTGG 2520  
2519 GCCTCAGGCACTCTTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2578  
2521 GCCTCAGGCACTCTTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2580

2579 ACAAGCTCTTCTCGACACCCCTACCTTTCTAATCAGGAGCAGTGGAGCAGTGCCT 2638  
2581 ACAAGCTCTTCTCGACACCCCTACCTTTCTAATCAGGAGCAGTGGAGCAGTGCCT 2640  
2639 CCTCTCTAGCACCTGCTTGTACCCAGCAAGAGGATAGGTTTGGAAACCTATCATTTCC 2698  
2641 CCTCTCTAGCACCTGCTTGTACCCAGCAAGAGGATAGGTTTGGAAACCTATCATTTCC 2700  
2699 TGTCTCTCTTAAAGAGGAAAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTTCG 2758  
2701 TGTCTCTCTTAAAGAGGAAAAAGCAGCTCTGTAGAAAGCAAGACCTTTCTTTTTCG 2760  
2759 GCTCTTTTCTTACAACTTAAAGCAGAAACTTCCAGAGTATTTGTTGGGGTGTGTT 2818  
2761 GCTCTTTTCTTACAACTTAAAGCAGAAACTTCCAGAGTATTTGTTGGGGTGTGTT 2820  
2819 TTATATTTAGGCAATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTCCTAA 2878  
2821 TTATATTTAGGCAATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTCCTAA 2880  
2879 GAAATGCTTAAACAAAGCACTTTTGGACAATGCTATCCAGCAGGAAAAAAAAGGATAA 2938  
2881 GAAATGCTTAAACAAAGCACTTTTGGACAATGCTATCCAGCAGGAAAAAAAAGGATAA 2940  
2939 TAACTGTTTAAACCTCTTTCTGGGGAATCCAAATTTATAGTTCTTGTATTTAAACAA 2998  
2941 TAACTGTTTAAACCTCTTTCTGGGGAATCCAAATTTATAGTTCTTGTATTTAAACAA 3000  
2999 GAAACAGCAAGGTTGTTGCGCAGGATAGGATGCTCTTAAAGATTTGGTCCCTTGAAT 3058  
3001 GAAACAGCAAGGTTGTTGCGCAGGATAGGATGCTCTTAAAGATTTGGTCCCTTGAAT 3060  
3059 ATGCTTCTGATCAAGGTAAGTATGTTGTCGCAACAAAGGAGGAACTTCTTTTAAAT 3118  
3061 ATGCTTCTGATCAAGGTAAGTATGTTGTCGCAACAAAGGAGGAACTTCTTTTAAAT 3120  
3119 TCCTTCTTCTTATTTTAAACAAATGTTGTAAGATGGAGGATTAATCAAAATCAGACAT 3178  
3121 TCCTTCTTCTTATTTTAAACAAATGTTGTAAGATGGAGGATTAATCAAAATCAGACAT 3180  
3179 GCAAAAACAAATGCTTGTGCTTCCATTAACAAAGTGAATTTTAAAGTGTCT 3238  
3181 GCAAAAACAAATGCTTGTGCTTCCATTAACAAAGTGAATTTTAAAGTGTCT 3240  
3239 TACTAAGTCTTGTATTAACCTCTCTTTATTTATATGGAATTAAGAGGAGGAGTCA 3298  
3241 TACTAAGTCTTGTATTAACCTCTCTTTATTTATATGGAATTAAGAGGAGGAGTCA 3300  
3299 TGTAGCAATGACAGCTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTCGATC 3358  
3301 TGTAGCAATGACAGCTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTCGATC 3360  
3359 CCTTCTGAGGATAGGCCCATCCAGACTTTTAGGCACTTTTGTAGGAAACAGATCCCTG 3418  
3361 CCTTCTGAGGATAGGCCCATCCAGACTTTTAGGCACTTTTGTAGGAAACAGATCCCTG 3420  
3419 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTAATACTGGATTTACATGAATC 3478  
3421 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTAATACTGGATTTACATGAATC 3480  
3479 TTTTGGTGTGTTCTATCAACCCACAGAGTTCCTAACTGCTTCAGTTTATAGTAAC 3538  
3481 TTTTGGTGTGTTCTATCAACCCACAGAGTTCCTAACTGCTTCAGTTTATAGTAAC 3540  
3539 TGACTGTGATATTTCAATCAGAGGCCCATAGCTAGTTGAGTATTTGATCCCTAGATAAG 3598  
3541 TGACTGTGATATTTCAATCAGAGGCCCATAGCTAGTTGAGTATTTGATCCCTAGATAAG 3600  
3599 AACATGCAAAATCAGCAGGAACTGCTCATAAGGGTAAGCAGGAGCAATAAGGATTTT 3658  
3601 AACATGCAAAATCAGCAGGAACTGCTCATAAGGGTAAGCAGGAGCAATAAGGATTTT 3660  
3659 TATAGATATAATTTAATTTTTTTGTTA - TTGTTAAGGAGA - CAATTTTGGAGAGCAGCA 3715

Db 3661 TATAGATATAATTTTAAATTTTGGTAAATGGGTAAAGGAGACCAATTTTGGAGGAAGCA 3720  
Qy 3716 AA---TCTTTTAAAAATAGTATGAAATGTAATGTAATAAGAAATTTAAAAAATAGTAT 3772  
Db 3721 AATCTTCTTTTAAAAAATAGTATGAAATGTAATGTAATAAGAAATTTAAGAAATAGTAT 3780  
Qy 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 3  
ABK83698  
ID ABK83698 standard; cDNA; 3802 BP.

XX AC ABK83698;  
XX 14-AUG-2002 (first entry)  
XX Human cDNA differentially expressed in granulocytic cells #269.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; periodontal disease;  
KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.  
XX WO200228999-A2.

XX 11-APR-2002.  
XX 03-OCT-2001; 2001WO-US030821.  
XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;  
XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression of  
PT genes associated with granulocyte activation, which serves as diagnostic  
PT markers that is useful for monitoring disease states and drug toxicity.

XX Claim 1; SEQ ID NO 269; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation  
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
CC DNA chip analysis as given in the specification, and comparing the  
CC expression level to an expression level in an unactivated GC, where  
CC differential expression of Gs is indicative of GCA. Also included are  
CC modulating (M2) GA by contacting GC with an agent that alters the  
CC expression of at least one gene in Gs; (2) screening (M3) for an agent  
CC capable of modulating GCA or an inflammation (especially chronic) in a  
CC tissue, an allergic response in a subject, exposure of a subject to a  
CC pathogen or sterile inflammatory disease using the gene expression  
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a  
CC tissue, an allergic response in a subject, exposure of a subject to a  
CC pathogen or sterile inflammatory disease, by detecting the level of  
CC expression in a sample of the tissue of gene(s) from Gs, where the level  
CC of expression of the gene is indicative of inflammation; (4) treating  
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic  
CC response in a subject, exposure of a subject to a pathogen or sterile  
CC inflammatory disease, by contacting a tissue having inflammation with an  
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1  
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful

CC for screening an agent capable of modulating GCA preferably in an  
CC inflammation in a tissue; M4 is useful for detecting an inflammation  
CC (especially chronic) in a tissue, an allergic response in a subject,  
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,  
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's  
CC disease, ulcerative colitis, periodontal disease; also bacterial  
CC infection, viral infection, parasitic infection, protozoal infection,  
CC fungal infection and M5 is useful for treating one of the above  
CC conditions. The present sequence represents a gene differentially  
CC expressed in granulocytes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 3802 BP; 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;

Query Match 98.1%; Score 3722; DB 6; Length 3802;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

Qy 1 ATAAATGACGTGCCGAGAGAGCGGAACCGCGAGCGGAGCGGAGTCTCTTGCT 60  
Db 1 ATAAATGACGTGCCGAGAGAGCGGAACCGCGAGCGGAGCGGAGTCTCTTGCT 60  
Qy 61 CCGCGCCCCCACCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 120  
Db 61 CCGCGCCCCCACCCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 120  
Qy 121 CACCGGCT 178  
Db 121 CACCGGCT 180  
Qy 179 ACTTCGCTCTCCGCGCGCTCACACCCCTCTTGCCTCAGCGCTTGCCTGCGGCGG 238  
Db 181 ACTTCGCTCTCCGCGCGCTCACACCCCTCTTGCCTCAGCGCTTGCCTGCGGCGG 240  
Qy 239 CGCGCAGCTGGAAGCGCCCTCCCGGGTCTCACTTTGCAACGCTGAGCGTGCAGGAGTGC 298  
Db 241 CGCGCAGCTGGAAGCGCCCTCCCGGGTCTCACTTTGCAACGCTGAGCGTGCAGGAGTGC 300  
Qy 299 CGTGAGGTGGGAAACAGCGCGCGCATCTCTCCCTCTGTGTACAGCCCAAGCCAGGACGCC 358  
Db 301 CGTGAGGTGGGAAACAGCGCGCGCATCTCTCCCTCTGTGTACAGCCCAAGCCAGGACGCC 360  
Qy 359 CGCGGAACCTCTCGGCTGTCTCTCCATGAGTGGGATCGCAGCATCCCCACCAGCG 418  
Db 361 CGCGGAACCTCTCGGCTGTCTCTCCATGAGTGGGATCGCAGCATCCCCACCAGCG 420  
Qy 419 CTCACCGGCTCCGGAGCGGCTGGGCTTGTATACCGCAGCCCTTCCGGGACAGCAGCTGT 478  
Db 421 CTCACCGGCTCCGGAGCGGCTGGGCTTGTATACCGCAGCCCTTCCGGGACAGCAGCTGT 480  
Qy 479 GACTCCCCCAGGTGCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGGAAC 538  
Db 481 GACTCCCCCAGGTGCAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGGAAC 540  
Qy 539 CGCCACAGCACTCAAAGCCCACTCGGAGAGGCGGAGCGCGCAAGCCCGGCGCTGAGC 598  
Db 541 CGCCACAGCACTCAAAGCCCACTCGGAGAGGCGGAGCGCGCAAGCCCGGCGCTGAGC 600  
Qy 599 CTGACACCTTAGCGGTGCGGGGACACTGCGGGGCTTGCCTCTCCCGGAGCTCCGCTC 658  
Db 601 CTGACACCTTAGCGGTGCGGGGACACTGCGGGGCTTGCCTCTCCCGGAGCTCCGCTC 660  
Qy 659 TTCTTACACTCTACGCTCTCGGTGAGAGAGACCCCGAGCCCACTTACGCGCGCAAGT 718  
Db 661 TTCTTACACTCTCTACGCTCTCGGTGAGAGAGACCCCGAGCCCACTTACGCGCGCAAGT 720  
Qy 719 ACCCTCCAGATATGCTCGGTCCAAAGCCCAATATAGCCCTTCCCTCTCCAGTTTCCAGTT 778  
Db 721 ACCCTCCAGATATGCTCGGTCCAAAGCCCAATATAGCCCTTCCCTCTCCAGTTTCCAGTT 780

QY 779 ATCCGGCGCAGATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACCA 838  
DB 781 ATCCGGCGCAGATACAGCTCGGAATACACACGAGATCATGAACCCCGACTACACCA 840  
QY 839 AGCTGACCATGACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGTCCCTGCCCA 898  
DB 841 AGCTGACCATGACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGTCCCTGCCCA 900  
QY 899 GCATCAGTACCTTCTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTGCGTGT 958  
DB 901 GCATCAGTACCTTCTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTGCGTGT 960  
QY 959 ACCAAATGACGGCCCTTGTATCAAGTGGAGAGGGCGGGCCAGCTACCATCACCC 1018  
DB 961 ACCAAATGACGGCCCTTGTATCAAGTGGAGAGGGCGGGCCAGCTACCATCACCC 1020  
QY 1019 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078  
DB 1021 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1080  
QY 1079 TTCTCTCAGCCTCCAGCCGAGGACGAGGTGCTGCCAGCACTCCATGTACTTCAAGC 1138  
DB 1081 TTCTCTCAGCCTCCAGCCGAGGACGAGGTGCTGCCAGCACTCCATGTACTTCAAGC 1140  
QY 1139 AGTCCCAACGCTCCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1198  
DB 1141 AGTCCCAACGCTCCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1200  
QY 1199 ACAGGACACTGCGCTCGCGCCGCGCTGCATCGCACCCGCGCGCTGCTGGACCCGCGCA 1258  
DB 1201 ACAGGACACTGCGCTCGCGCCGCGCTGCATCGCACCCGCGCGCTGCTGGACCCGCGCA 1260  
QY 1259 TGAAGCGGTGCCAGCTGCGCGCGCGCTGCTGCCAGCACTTCAAGCCCTCGC 1318  
DB 1261 TGAAGCGGTGCCAGCTGCGCGCGCGCTGCTGCCAGCACTTCAAGCCCTCGC 1320  
QY 1319 CCGCGCATCCCGCGCCAGCCCG 1378  
DB 1321 CCGCGCATCCCGCGCGCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1379 CCGCTGCGCGCTCAGCTGCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1438  
DB 1381 CCGCTGCGCGCTCAGCTGCGCGCTGGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
QY 1439 CCGCTTGAAGCACCACCGTACGGGCTGCGCGCTGCGCAAGGCGCGCGCGCGCGCGCGCGCG 1498  
DB 1441 CCGCTTGAAGGCGCACCGGTACGGGCTGCGCGCTGCGCAAGGCGCGCGCGCGCGCGCGCG 1500  
QY 1499 CCGCTCTCGGCTCAGCGCTCCCTTACCGGCTCGAGCTGCTGGCGGAGAGTCCAGCC 1558  
DB 1501 CCGCTCTCGGCTCAGCGCTCCCTTACCGGCTCGAGCTGCTGGCGGAGAGTCCAGCC 1560  
QY 1559 TCGCGTGCAGCGCGCGCGCGCTGCTGCTGGCGAGGCGCGCTGCGCGTGGCGG 1618  
DB 1561 TCGCGTGCAGCGCGCGCGCGCGCTGCTGCTGGCGAGGCGCGCTGCGCGTGGCGG 1620  
QY 1619 ACAACGCGCGCTGCCAGCACTACGGGCTGCGAACCTGCGAGGCGTGCAGAGGCTTTTCA 1678  
DB 1621 ACAACGCGCGCTGCCAGCACTACGGGCTGCGAACCTGCGAGGCGTGCAGAGGCTTTTCA 1680  
QY 1679 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAACTGCCAGTAG 1738  
DB 1681 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCGCAATAAAACTGCCAGTAG 1740  
QY 1739 ACAAGAGAGTGCAGAACCTGCTAGTACTGCTGATTTTCAAGAGTGTCTCAGTGTGGAA 1798  
DB 1741 ACAAGAGAGTGCAGAACCTGCTAGTACTGCTGATTTTCAAGAGTGTCTCAGTGTGGAA 1800  
QY 1799 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGAGGAGGAGAGTGTGCTGCTGCCCTTCA 1858  
DB 1801 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGAGGAGGAGAGTGTGCTGCTGCCCTTCA 1860

QY 1859 AACCAAGAGCCCATTAACAAGGACCTTCTCAGCCCTCTCCACCTTCTCTCCCAATCT 1918  
DB 1861 AACCAAGAGCCCATTAACAAGGACCTTCTCAGCCCTCTCCACCTTCTCTCCCAATCT 1920  
QY 1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTA 1978  
DB 1921 GCATGATGAATGCTCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTA 1980  
QY 1979 CCAGATACGTGTCCTCCACTGACACAGGCTGCTGAGGCAACAGATGCTGAGCAACAAT 2038  
DB 1981 CCAGATACGTGTCCTCCACTGACACAGGCTGCTGAGGCAACAGATGCTGAGCAACAAT 2040  
QY 2039 TCTACAACTCTCTGACAGCTCCATGTATCCAGAGCTGGGAGGAGAAAGATTTCCG 2098  
DB 2041 TCTACAACTCTCTGACAGCTCCATGTATCCAGAGCTGGGAGGAGAAAGATTTCCG 2100  
QY 2099 GATTTACTGATCTCCCAAGAGAGATCAGACATTAATTAATCAGCCTTTTGGAGC 2158  
DB 2101 GATTTACTGATCTCCCAAGAGAGATCAGACATTAATTAATCAGCCTTTTGGAGC 2160  
QY 2159 TGTGTGCTCAGACTTTCCATCAGGTCAACACACTGCTGAAGATAAGTTTGTCTGCA 2218  
DB 2161 TGTGTGCTCAGACTTTCCATCAGGTCAACACACTGCTGAAGATAAGTTTGTCTGCA 2220  
QY 2219 ATGGACTGTCTGCTGATCAGCTTCTGAGCTTCTGAGATTTGGGAGTGGCTGCACTTA 2278  
DB 2221 ATGGACTGTCTGCTGATCAGCTTCTGAGCTTCTGAGATTTGGGAGTGGCTGCACTTA 2280  
QY 2279 TTAAGACCTTTTCTTAAATTTGAGAGCTCAACCTTGATATCAAGCCTTAGCCTGCC 2338  
DB 2281 TTAAGACCTTTTCTTAAATTTGAGAGCTCAACCTTGATATCAAGCCTTAGCCTGCC 2340  
QY 2339 TGTGAGCACTGAGCATGATCAGAAAGACATGGGTAAAAAGACCAAGAGAGTCAAG 2398  
DB 2341 TGTGAGCACTGAGCATGATCAGAAAGACATGGGTAAAAAGACCAAGAGAGTCAAG 2400  
QY 2399 AGCTATGCAACAGATCAACAGCAGTTTAAAGACCAACAGAGTAAGGACAGCTCTG 2458  
DB 2401 AGCTATGCAACAGATCAACAGCAGTTTAAAGACCAACAGAGTAAGGACAGCTCTG 2460  
QY 2459 AGCCACACGAGTCAAGGTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2518  
DB 2461 AACCCAAAGAGTCAAGGTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2520  
QY 2519 GCCTCCAGCGCATCTTCTACCTGAAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCTG 2578  
DB 2521 GCCTCCAGCGCATCTTCTACCTGAAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCTG 2580  
QY 2579 ACAAGCTCTTCTGAGACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGGCT 2638  
DB 2581 ACAAGCTCTTCTGAGACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGGCT 2640  
QY 2639 CCTCTCTAGCACCTTGTCTGCTACGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2698  
DB 2641 CCTCTCTAGCACCTTGTCTTACGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2700  
QY 2699 TGTCTCTCTTAAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGAGCTTTCTTTTCTG 2758  
DB 2701 TGTCTCTCTTAAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGAGCTTTCTTTTCTG 2760  
QY 2759 GCTCTTTCTCTTAACCTAAAGCAGAAACCTGAGAGTATGTTGTTGGGTTGTGTT 2818  
DB 2761 GCTCTTTCTCTTAACCTAAAGCAGAAACCTGAGAGTATGTTGTTGGGTTGTGTT 2820  
QY 2819 TTATATTTAGGATTTGGGATTTGGGATTTGGGATTTAGTTTATAGTTTATTTTAA 2878  
DB 2821 TTATATTTAGGATTTGGGATTTGGGATTTGGGATTTAGTTTATAGTTTATTTCTAA 2880  
QY 2879 GAAATTTGCTTAAACAAAGCCTTTTGAACAATGCTATCCAGCAGGAAAAAGGATAA 2938  
DB 2881 GAAATTTGCTTAAACAAAGCCTTTTGAACAATGCTATCCAGCAGGAAAAAGGATAA 2940  
QY 2939 TAACTGTTTAAACACTCTTTCTGGGGAATCCAATTAATAGTTGCTTTGTTAAAAACAA 2998



CC proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4 (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a shows homology to matrix metalloproteinase-15 precursor; NOV4a shows homology to MD3; NOV5a-5c show homology to T-lymphocyte surface antigen ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21 (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like CC polypeptide variant; NOV8a-8g show homology to antileukoprotease 1 CC precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology CC to nuclear hormone receptor NOR-1; NOV11a-11j show homology to CC transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin CC -dynorphin precursor. The present sequence encodes a NOVX polypeptide of CC the invention.

XX Sequence 3802 BP; 919 A; 1141 C; 912 G; 830 T; 0 U; 0 Other;

SQ Query Match 98.1%; Score 3722; DB 12; Length 3802;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 1 ATAAATGAGTGTCCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGTCTCTGCGCT 60  
DB 1 ATAAATGAGTGTCCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGTCTCTGCGCT 60  
QY 61 CCGCGCCCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCCATACACAGCGGCTCA 120  
DB 61 CCGCGCCCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCCATACACAGCGGCTCA 120  
QY 121 CACCGGCTCCCTCACTCGCACACACAGACAGCGGCGCACAGGCTCCG--CACACAC 178  
DB 121 CACCGGCTCCCTCACTCGAACACACAGACAGCGGCGCACAGGCTCCGCGACACAC 180  
QY 179 ACTTCGCTCTCCGGCGGCTCACACCCCTTTGGCTTGGCTTGGCTTGGCTTGGCT 238  
DB 181 ACTTCGCTCTCCGGCGGCTCACACCCCTTTGGCTTGGCTTGGCTTGGCTTGGCT 240  
QY 239 CGCGGAGTGGAGCGGCTCCGCGGCTCACTTTGCAACGCTGACGCTCCGCGAGTGGC 298  
DB 241 CGCGGAGTGGAGCGGCTCCGCGGCTCACTTTGCAACGCTGACGCTCCGCGAGTGGC 300  
QY 299 CGTGGAGTGGGAACAGCGGCGGATCTCCCTCCGCTTGGTACAGCCCAAGCCAGGACGC 358  
DB 301 CGTGGAGTGGGAACAGCGGCGGATCTCCCTCCGCTTGGTACAGCCCAAGCCAGGACGC 360  
QY 359 CGCGGAACCTCTCGGCTGTCTCTCCATGAGTCGGGATCGAGCATCCCGACAGCGG 418  
DB 361 CGCGGAACCTCTCGGCTGTCTCTCCATGAGTCGGGATCGAGCATCCCGACAGCGG 420  
QY 419 CTCACCGCTCCGGGAGCGCTGGGCTGTACACCGCAGCCCTTCCGGGACGAGCTGT 478  
DB 421 CTCACCGCTCCGGGAGCGCTGGGCTGTACACCGCAGCCCTTCCGGGACGAGCTGT 480  
QY 479 GACTCCCGCCAGTGCAGATTTGGGACAGCTCTCTAGAAAACCTCGCTCTAAAGACGGAAC 538  
DB 481 GACTCCCGCCAGTGCAGATTTGGGACAGCTCTCTAGAAAACCTCGCTCTAAAGACGGAAC 540  
QY 539 CGGCAAGCACTCAAAAGCCACTGCGGAAAGAGGCGAGCCCGGAGCCCGGCTTGGAGC 598  
DB 541 CGGCAAGCACTCAAAAGCCACTGCGGAAAGAGGCGAGCCCGGAGCCCGGCTTGGAGC 600  
QY 599 CTGGACCCCTTAGGGTGGCGGAGCAGCTCCGCGGCTTCCGCTCCGCGGAGCTCCGCTC 658  
DB 601 CTGGACCCCTTAGGGTGGCGGAGCAGCTCCGCGGCTTCCGCTCCGCGGAGCTCCGCTC 660  
QY 659 CTCCTACACTCTCAGGCTCGGTGGAGAGACCCCGAGCCCGGCTTCCGCTCCGCGGAGAT 718  
DB 661 CTCCTACACTCTCAGGCTCGGTGGAGAGACCCCGAGCCCGGCTTCCGCTCCGCGGAGAT 720  
QY 719 ACCCTCCAGATATGCCCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 778  
DB 721 ACCCTCCAGATATGCCCTCGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 780  
QY 779 ATGGCGGCGAGACATACGCTGGGAATACACACGAGAGATCATGAACCCCGGACTACCA 838  
DB ATGGCGGCGAGACATACGCTGGGAATACACACGAGAGATCATGAACCCCGGACTACCA

DB 781 ATGGCGGCGAGACATACAGCTCGGAATACACACGAGAGATCATGAACCCCGGACTACACCA 840  
QY 839 AGCTGACCAATGAGACCTTGGCAGACCTAGATCAAGCTTACAGCCACCAAGTCTCTGCCCA 898  
DB 841 AGCTGACCAATGAGACCTTGGCAGACCTAGATCAAGCTTACAGCCACCAAGTCTCTGCCCA 900  
QY 899 GCATCAGTACCTTCCGTTGGAGGCTTCTCGAGCAATACCAACTCAAGGCTTCTCGGCTGT 958  
DB 901 GCATCAGTACCTTCCGTTGGAGGCTTCTCGAGCAATACCAACTCAAGGCTTCTCGGCTGT 960  
QY 959 ACCAAATGAGCGGCGCTTGGATCAAAAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1018  
DB 961 ACCAAATGAGCGGCGCTTGGATCAAAAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1020  
QY 1019 ATCAAC 1078  
DB 1021 ATCAAC 1080  
QY 1079 TTTCTCCAGCTTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1138  
DB 1081 TTTCTCCAGCTTCCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
QY 1139 AGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1198  
DB 1141 AGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200  
QY 1199 ACGAGGAGTGTCCGCTTCCGCGGCGGCTGATCGACACCGGCGGCGGCTGTGGAACCCCGG 1258  
DB 1201 ACGAGGAGTGTCCGCTTCCGCGGCGGCTGATCGACACCGGCGGCGGCTGTGGAACCCCGG 1260  
QY 1259 TGAAGGCGGTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1318  
DB 1261 TGAAGGCGGTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1319 CGCGCATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1378  
DB 1321 CGCGCGATCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380  
QY 1379 CGGCTGCGGCTTCAAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1438  
DB 1381 CGGCTGCGGCTTCAAGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1440  
QY 1439 CGCTTGAGAGCACCGGCTACGGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1498  
DB 1441 CGCTTGAGGCGCACCGTACGGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1500  
QY 1499 CGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1558  
DB 1501 CGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1560  
QY 1559 TGCCGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1618  
DB 1561 TGCCGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620  
QY 1619 ACAACGCGGCTTCCGAGCACTACGGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 1678  
DB 1621 ACAACGCGGCTTCCGAGCACTACGGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 1680  
QY 1679 AGAGAACAGTGCAGAAATATGTTTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 1738  
DB 1681 AGAGAACAGTGCAGAAATATGTTTGGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 1740  
QY 1739 ACAAGAGAGCTCGAAACCGATGTCTAGTACTGTCTGATTTTCCGCTTCCGCTTCCGCTT 1798  
DB 1741 ACAAGAGAGCTCGAAACCGATGTCTAGTACTGTCTGATTTTCCGCTTCCGCTTCCGCTT 1800  
QY 1799 TGGTAAAGAGAGTGTCTCGGTACAGATAGTCTGAAAGGAGGAGAGGAGGAGGAGGAGGAG 1858  
DB 1801 TGGTAAAGAGAGTGTCTCGGTACAGATAGTCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1860  
QY 1859 AACCAAGAGGCGGCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1918  
DB 1861 AACCAAGAGGCGGCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920

```
QY 1919 GCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTATT 1978
Db 1921 GCATGATGAATGCTCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGATTATT 1980
QY 1979 CCAGATACCTGCTCCCACTGACAGAGCTGCTGAGGACACAGATGCTGAGCATGTGCAACAAT 2038
Db 1981 CCAGATACCTGCTCCCACTGACAGAGCTGCTGAGGACACAGATGCTGAGCATGTGCAACAAT 2040
QY 2039 TCTACAACTCTGACAGCTCCATTGATGATATCAGAGCTGCGACGAAAGATTCCGG 2098
Db 2041 TCTACAACTCTGACAGCTCCATTGATGATATCAGAGCTGCGACGAAAGATTCCGG 2100
QY 2099 GATTTACTGATCTCCCAAGAGATCAGACATTAATTATTGAATCAGCTTTTGGAGC 2158
Db 2101 GATTTACTGATCTCCCAAGAGATCAGACATTAATTATTGAATCAGCTTTTGGAGC 2160
QY 2159 TGTGTGCTCAGACTTTCCATCAGGTCAACACACTGCTGAAGATTAAGTTGTCTGCA 2218
Db 2161 TGTGTGCTCAGACTTTCCATCAGGTCAACACACTGCTGAAGATTAAGTTGTCTGCA 2220
QY 2219 ATGGACTTGTCTGATCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2278
Db 2221 ATGGACTTGTCTGATCGACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTA 2280
QY 2279 TTAAGAGACTTTTCTTAAATTTGCAGAGCTGAACCTTGATATCAAGCCTTAGCCTGCC 2338
Db 2281 TTAAGAGACTTTTCTTAAATTTGCAGAGCTGAACCTTGATATCAAGCCTTAGCCTGCC 2340
QY 2339 TGTGACGATGAGCATGATCAGAGAAAGACATGGGTAAAGAACCAAGAGAGTCGAAG 2398
Db 2341 TGTGACGATGAGCATGATCAGAGAAAGACATGGGTAAAGAACCAAGAGAGTCGAAG 2400
QY 2399 AGCTATGAAACAAAGATCAACAGCAGTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2458
Db 2401 AGCTATGAAACAAAGATCAACAGCAGTTAAAGACACACAGAGTAAGGACAGGCTCTGG 2460
QY 2459 AGCCCAACGAGTCCAAGGTCCTGGGTGCTTAAAGACCTGAGAGATCTGACCCCTGG 2518
Db 2461 AACCACCAAGTCCAAGGTCCTGGGTGCTTAAAGACCTGAGAGATCTGACCCCTGG 2520
QY 2519 GCCTCAGGCACTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTG 2578
Db 2521 GCCTCAGGCACTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTG 2580
QY 2579 ACAAGCTCTTCTGACACCTTCTACCTTCTAATCAGGAGCAGTGAGCAGTGCCT 2638
Db 2581 ACAAGCTCTTCTGACACCTTCTACCTTCTAATCAGGAGCAGTGAGCAGTGCCT 2640
QY 2639 CCTCTCCTAGCACCTGCTGTACGACGAAAGGATAGGTTTGGAAACCTTATCATTTCC 2698
Db 2641 CCTCTCCTAGCACCTGCTGTACGACGAAAGGATAGGTTTGGAAACCTTATCATTTCC 2700
QY 2699 TGCTCTCTTAAAGAGAAAGCAGCTCTGTAGAGCAAGACTTCTTTTTCGT 2758
Db 2701 TGCTCTCTTAAAGAGAAAGCAGCTCTGTAGAGCAAGACTTCTTTTTCGT 2760
QY 2759 GCTCTTTTCTTACAACTTAAAGCAGAAACTGTCAGAGTATGTGTGGGTTGTGT 2818
Db 2761 GCTCTTTTCTTACAACTTAAAGCAGAAACTTGCAGAGTATGTGTGGGTTGTGT 2820
QY 2819 TTATATTTAGGCATTTGGGGATTTGGGATGAGGGGTTATAGTTTATGAGGGTTTCTAA 2878
Db 2821 TTATATTTAGGCATTTGGGGATTTGGGATGAGGGGTTATAGTTTATGAGGGTTTCTAA 2880
QY 2879 GAAATGCTTAAACAGCACTTTGGACATGCTATCCAGCAGGAAAGGATTAATA 2938
Db 2881 GAAATGCTTAAACAGCACTTTGGACATGCTATCCAGCAGGAAAGGATTAATA 2940
QY 2939 TAACTGTTTTAAACCTTTCTCTGGGAAATCCAAATTATAGTTTGTATTTAAACAA 2998
Db 2941 TAACTGTTTTAAACCTTTCTCTGGGAAATCCAAATTATAGTTTGTATTTAAACAA 3000
```

```
QY 2999 GAACAGCAAGGGTTGTTCCGAGGTAGGATGTCTTAAAGATTGTCCTCTGAAAT 3058
Db 3001 GAACAGCAAGGGTTGTTCCGAGGTAGGATGTCTTAAAGATTGTCCTCTGAAAT 3060
QY 3059 ATGCTTCTGTTATCAAGGTACGTATGTGTGCAAAACAAAGCAGAACTTCTTTTAAAT 3118
Db 3061 ATGCTTCTGTTATCAAGGTACGTATGTGTGCAAAACAAAGCAGAACTTCTTTTAAAT 3120
QY 3119 TCCTTCTTCTTTTAAACAAATGTTGAAGATTGAGGATTACCTTACAAATCAGACAT 3178
Db 3121 TCCTTCTTCTTTTAAACAAATGTTGAAGATTGAGGATTACCTTACAAATCAGACAT 3180
QY 3179 GGCATAAACAATTAATGGCTTTTGTTCATATAACAAAGTGCATTTTAAAGTGTCT 3238
Db 3181 GGCATAAACAATTAATGGCTTTTGTTCATATAACAAAGTGCATTTTAAAGTGTCT 3240
QY 3239 TACTAAGTCTTGTATTATTAACCTCTCTTATTTCTATATGGAATAAAGAGGAGGCA 3298
Db 3241 TACTAAGTCTTGTATTATTAACCTCTCTTATTTCTATATGGAATAAAGAGGAGGCA 3300
QY 3299 TGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGATC 3358
Db 3301 TGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGATC 3360
QY 3359 CCTTCTGAGGTATGGCCCATCCAAAGCTTTTGTAGGCCATTTGTATGGAACAGATCCCTG 3418
Db 3361 CCTTCTGAGGTATGGCCCATCCAAAGCTTTTGTAGGCCATTTGTATGGAACAGATCCCTG 3420
QY 3419 CCCTGACGTCCGACTATCCCTGAAAGTGGATCAGATATTAACCTGATTACATGTAAGCTG 3478
Db 3421 CCCTGACGTCCGACTATCCCTGAAAGTGGATCAGATATTAACCTGATTACATGTAAGCTG 3480
QY 3479 TTTTGGTTGTCTTCTATCAACCCCAACAGAGTTCCTTAAACTTGTCTCAGTTATAGTAAC 3538
Db 3481 TTTTGGTTGTCTTCTATCAACCCCAACAGAGTTCCTTAAACTTGTCTCAGTTATAGTAAC 3540
QY 3539 TGACTGTGATATTCATTTCAGAGGCCCATAGTCAGTGTGATTTGATCCCTAGATAAG 3598
Db 3541 TGACTGTGATATTCATTTCAGAGGCCCATAGTTCAGTGTGATTTGATCCCTAGATAAG 3600
QY 3599 AACATGCAAAATCAGCAGGAACTGTGTATACAGGGTAAAGCAGGACCAATAAGGATTTT 3658
Db 3601 AACATGCAAAATCAGCAGGAACTGTGTATACAGGGTAAAGCAGGACCAATAAGGATTTT 3660
QY 3659 TATAGATATAAATTAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3715
Db 3661 TATAGATATAAATTAATTTTGTTA--TTGGTTAAGGAGACCAATTTTGGAGAGCAAGCA 3720
QY 3716 AA---TCCTTTTAAAAATAGTATGAATGTGAATACTAGAAAAAGATTTTAAAAATAGTAT 3772
Db 3721 AATCTTCTTTTAAAAATAGTATGAATGTGAATACTAGAAAAAGATTTTAAAAATAGTAT 3772
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802
```

## RESULT 5

ADE25701

ID ADE25701 standard; cDNA; 5828 BP.

XX ADE25701;

AC ADE25701;

XX 29-JAN-2004 (first entry)

XX Human cDNA differentially expressed in foam cells #105.

DE Human; ss; differential expression; foam cell; lipopolysaccharide;

KW cardiovascular disease; atherosclerosis.

XX Homo sapiens.

XX OS

XX US2003194721-A1.









3541	ACTGACTGGTATATTCATTCAGAAAGCCCAATAGTCACTAGTTGAGTATTTGTGATTCCTAGATA	3600
Db		
3597	AGAAATCATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACACAGGGAACAATAAGGATT	3656
Qy		
3601	AGAAATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACACAGGGAACAATAAGGATT	3660
Db		
3657	TTTATAGATATAATTTTAAATTTTGTATTTCGGTTAAAGGAGACAATTTTGGAGAGCAAGCAA	3716
Qy		
3661	TTTATAGATATAATTTTAAATTTTGTATTTCGGTTAAAGGAGACAATTTTGGAGAGCAAGCAA	3720
Db		
3717	A---TCCTTTTAAAAAATAGTATGCAATGTCGAATACTAGAAAAGATTTTAAAAAATAGTATG	3773
Qy		
3721	ATCTTCTTTTTAAAAAATAGTATGAAATGTGAATACTAGAAAAGATTTTAAGAAATAGTATG	3780
Db		
3774	AGTGTGAGTACTAGGAAGGAT	3794
Qy		
3781	AGTGTGAGTACTAGGAAGGAT	3801
Db		

## RESULT 6

ABX62937	
ID	ABX62937 standard; cDNA; 5642 BP.
XX	
XX	AC
XX	ABX62937;
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human activated T cell cDNA #53.
XX	
XX	T cell; gene; ss; differential expression; T cell activation;
KW	antiallergic; cytostatic; immunosuppressive; antimicrobial; gene therapy;
KW	allergy; cancer; graft versus host disease; infection;
KW	autoimmune disorder.
XX	
OS	Homo sapiens.
XX	
PN	US2002137077-A1.
XX	
PD	26-SEP-2002.
XX	
PF	25-OCT-2001; 2001US-0002600.
XX	
PR	25-OCT-2000; 2000US-0243521P.
XX	
PA	(HOPK/) HOPKINS C M.
PA	(PETE/) PETERSON D P.
PA	(COCK/) COCKS B G.
PA	(HAWK/) HAWKINS P R.
XX	
PI	Hopkins CM, Peterson DP, Cocks BG, Hawkins PR;
XX	
WPI	2003-102381/09.
XX	
PT	New combination comprising several cDNAs that are differentially
PT	expressed in activated T cells, useful for diagnosing, treating, staging
PT	or monitoring treatment for allergy, cancer, infectious and/or autoimmune
PT	disorders.
XX	
PS	Claim 1; Page; 180pp; English.
XX	
CC	This invention relates to the sequences of several cDNAs that are
CC	differentially expressed in activated T cells. The sequences of the
CC	invention may have antiallergic, cytostatic, immunosuppressive and
CC	antimicrobial activity and may be used in gene therapy. The invention
CC	also comprises a method for screening samples for differentially
CC	expressed genes and a method for detecting these cDNAs by hybridisation.
CC	The methods and compositions of the present invention are useful for,
CC	diagnosing, treating, staging or monitoring treatment for allergy,
CC	cancer, chronic graft versus host disease, infectious and/or autoimmune
CC	disorders. The present sequence represents a cDNA of the invention that
CC	is differentially expressed in activated T cells
XX	
SQ	Sequence 5642 BP; 1472 A; 1399 C; 1212 G; 1465 T; 0 U; 94 Other;



```

Db 3181 CATGCCAAACAATAATGCTGTTTCTTCCATAAACAAAGTGCAATTTTTTAAAGTGCTG 3240
QY 3236 TCTTACTAAGTCTTGTTTATTAACCTCTCTTTATCTATATGGAATAAAGAGGCGAG 3295
Db 3241 TCTTACTAAGTCTTGTTTATTAACCTCTCTTTATCTATATGGAATAAAGAGGCGAG 3300
QY 3296 TCATGTTAGCAATGACAGCTTAATATCCCTAGCAGAGCTGTGTTTACCTTCCCTGTGCG 3355
Db 3301 TCATGTTAGCAATGACAGCTTAATATCCCTAGCAGAGCTGTGTTTACCTTCCCTGTGCG 3360
QY 3356 ATCCCTCTTGAGGTATGGCCCATCCAAGACTTTTAGGCCAATCTTGTATGGAACCATGATCC 3415
Db 3361 ATCCCTCTTGAGGTATGGCCCATCCAAGACTTTTAGGCCAATCTTGTATGGAACCATGATCC 3420
QY 3416 CTGCCCTGACTGTCAGCTATCTGAAAGTGGATCAGATATATAAAGTGAATACATGTA 3475
Db 3421 CTGCCCTGACTGTCAGCTATCTGAAAGTGGATCAGATATATAAAGTGAATACATGTA 3480
QY 3476 CTGTTTGGTGTGTTTCTATCAACCCACCAGAGTTCCCTAAACTTGTCTCAGTTATAGT 3535
Db 3481 CTGTTTGGTGTGTTTCTATCAACCCACCAGAGTTCCCTAAACTTGTCTCAGTTATAGT 3540
QY 3536 AACTGACTGTTATATTCATTTCAGAGCGCCATAAGTTCAGTTGATTTGATCCCTAGAT 3595
Db 3541 AACTGACTGTTATATTCATTTCAGAGCGCCATAAGTTCAGTTGATTTGATCCCTAGAT 3600
QY 3596 AAGAACATCAATCAGCAGGAACCTGTCATCAGGGTAAGCACAAGGAGCAATAGGAT 3655
Db 3601 AAGAACATCAATCAGCAGGAACCTGTCATCAGGGTAAGCACAAGGAGCAATAGGAT 3660
QY 3656 TTTTATAGATATAATTTAATTTTGTATTGTTTAAAGGAGCAATTTTGGAGAGAAGCA 3715
Db 3661 TTTTATAGATATAATTTAATTTTGTATTGTTTAAAGGAGCAATTTTGGAGAGAAGCA 3720
QY 3716 AA---TCTTTTAAATAATAGTATGAATGTGAATCTAGAAAGAGTTTAAATAATAGTAT 3772
Db 3721 AATCTCTTTTAAATAATAGTATGAATGTGAATCTAGAAAGAGTTTAAATAATAGTAT 3780
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

```

## RESULT 7

AAT73334  
ID AAT73334 standard; cDNA to mRNA; 3762 BP.

AC AAT73334;

XX 08-SEP-1997 (first entry)

XX Human neuron-derived orphan receptor NOR-1 gene.

DE Human; neuron-derived orphan receptor; open reading frame; homology; rat;  
KW NOR-1; DNA binding domain; brain; neuronal; Alzheimer's disease;  
KW Parkinson's disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

PH 692..2572

FT /\*tag= a

FT /product= "NOR-1 orphan receptor"

XX JP09084585-A.

XX 31-MAR-1997.

XX 21-SEP-1995; 95JP-00242909.

XX 21-SEP-1995; 95JP-00242909.

XX (TERU ) TERUMO CORP.

XX WPI; 1997-253003/23.  
DR P-PSDB; AAW16398.  
XX Human neuron-derived orphan nuclear receptor gene - useful for diagnosing  
PT brain diseases such as Alzheimer's or Parkinson's disease or  
PT neuroblastoma.  
XX Claim 1; Page 5-8; 8pp; Japanese.  
XX This is the nucleotide sequence of a novel human neuron-derived orphan  
CC nuclear receptor gene. The open reading frame encodes a protein of 626  
CC amino acids, which has high homology to the rat NOR-1 receptor. The  
CC protein contains several DNA binding domains including P, D and A boxes  
CC which also have high homology to the rat NOR-1 DNA binding domains. The  
CC nucleic acid and protein can be used to diagnose brain neuronal diseases,  
CC such as Alzheimer's disease, Parkinson's disease. Further, the genetic  
CC diseases can be treated using the receptor as it inhibits expression of a  
CC gene by binding to it  
XX Sequence 3762 BP; 907 A; 1128 C; 898 G; 829 T; 0 U; 0 Other;

Query Match 96.3%; Score 3653.6; DB 2; Length 3762;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 3728; Conservative 0; Mismatches 24; Indels 10; Gaps 5;

QY 43 GAGCGGAGTCTCTGCTCCGCGCCGCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCC 102

Db 1 GAGCGGAGTCTCTGCTCCGCGCCGCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCC 60

QY 103 CATACACAGACGCGCTCACACCGCTCCTCTCACTCGCACACACAGACACAAGCGGCACA 162

Db 61 CATACACAGACGCGCTCACACCGCTCCTCTCACTCGCACACACAGACACAAGCGGCACA 120

QY 163 CAGGCTCCG--CACACACACTTTCGCTCTCCGCGCGCTCACACCCCTCTTGGCTGAGGC 220

Db 121 CAGGCTCCGACACACACACTTTCGCTCTCCGCGCGCTCACACCCCTCTTGGCTGAGGC 180

QY 221 CTTGCCGCTGACGCGCGCGCGAGCTGGAGCCCTCCCGGGTCACTTTGCAACGCT 280

Db 181 CTTGCCGCTGACGCGCGCGCGAGCTGGAGCCCTCCCGGGTCACTTTGCAACGCT 240

QY 281 GACGCTGCGGAGTGGCGCTGGAGTGGGAACACGCGGGGATCTCTCCCTCTGGTAC 340

Db 241 GACGCTGCGGAGTGGCGCTGGAGTGGGAACACGCGGGGATCTCTCCCTCTGGTAC 300

QY 341 AGCCCAA--GCCAGGACGCCCGCGAACTCTCGGCTGTGCTCTCCATGAGTCCGGATC 398

Db 301 AGCCCAAAGCCAGGAGCGCCCGGAACTCTCGGCTGTGCTCTCCATGAGTCCGGATC 360

QY 399 GCAGCATCCCCACACGCGCTCAACGCTCCGGAGCGCTGGGCTTGTATACCGGACG 458

Db 361 GCAGCATCCCCACACGCGCTCAACGCTCCGGAGCGCTGGGCTTGTATACCGGACG 420

QY 459 CTTCCGGGACAGAGCTGTGACTCCCGCCAGTCAGATTTCGGGACAGCTCTCTAGAA 518

Db 421 CTTCCGGGACAGAGCTGTGACTCCCGCCAGTCAGATTTCGGGACAGCTCTCTAGAA 480

QY 519 ACTGCTCTTAAAGACGGAACCGCCACAGCACTCAAAGCCCACTGCGGAAGAGGAGCCCC 578

Db 481 ACTGCTCTTAAAGACGGAACCGCCACAGCACTCAAAGCCCACTGCGGAAGAGGAGCCCC 540

QY 579 GGCAAGCCCGGCGCTGAGCTTGGACCTTAGCGGTGCGGGACAGCTCTCCGGGCTTC 638

Db 541 GGCAAGCCCGGCGCTGAGCTTGGACCTTAGCGGTGCGGGACAGCTCTCCGGGCTTC 600

QY 639 GCCTCGCGGAGTCCGCTCTCTTACACTCTCAGCTCCGCTCGGTGAGAGACCCCGCCCC 698

Db 601 GCCTCGCGGAGTCCGCTCTCTTACACTCTCAGCTCCGCTCGGTGAGAGACCCCGCCCC 660

QY 699 CACCATTCAGCGCGCAAGATACCTCCAGATATGCCCTCCAGTCCAGCCCATATAGCCC 758

Db 661 CACCATTCAGCGCGCAAGATACCTCCAGATATGCCCTCCAGTCCAGCCCATATAGCCC 720

QY 759 TTCCCTCCAGTTCAGTTATGCGCGCAGACATACAGCTCGGAATACACCGAGAT 818  
Db 721 TTCCCTCCAGTTCAGTTATGCGCGCAGACATACAGCTCGGAATACACCGAGAT 780  
QY 819 CATGAACCCCGACTACACCAAGCTGACCATGGACCTTGGCGACCTGAGATCAAGGCTAC 878  
Db 781 CATGAACCCCGACTACACCAAGCTGACCATGGACCTTGGCGACCTGAGATCAAGGCTAC 840  
QY 879 AGCCACACGTCCTGCGCGCAGCATCAGTACCTTGGTGGAGGCTACTCGAGCAACTACGA 938  
Db 841 AGCCACACGTCCTGCGCGCAGCATCAGTACCTTGGTGGAGGCTACTCGAGCAACTACGA 900  
QY 939 ACTCAAGCCTTCTCGCTGTAACCAATGACGGCCCTTTGATCAAGTGGAGGAGGGCG 998  
Db 901 ACTCAAGCCTTCTCGCTGTAACCAATGACGGCCCTTTGATCAAGTGGAGGAGGGCG 960  
QY 999 GCGGCCAGCTACCATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1058  
Db 961 GCGGCCAGCTACCATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1020  
QY 1059 GCAGCATCAGACGCTCATCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1118  
Db 1021 GCAGCATCAGACGCTCATCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1080  
QY 1119 CACCTCCATGTACTTCAAGCAGTCCCAACCGTCCCAACCGTCCCAACCGTCCCAACCGT 1178  
Db 1081 CACCTCCATGTACTTCAAGCAGTCCCAACCGTCCCAACCGTCCCAACCGTCCCAACCGT 1140  
QY 1179 CGAGCGGGGGGTTATGGAGAGGCACTGCGCTCGCGCGCGCGCTGCATCGCACCGG 1238  
Db 1141 CGAGCGGGGGGTTATGGAGAGGCACTGCGCTCGCGCGCGCGCTGCATCGCACCGG 1200  
QY 1239 CCGGTGCTGACCCGCGCGTGAAGCGCGTCCCAACCGTCCCAACCGTCCCAACCGTCCCGCT 1298  
Db 1201 CCGGTGCTGACCCGCGCGTGAAGCGCGTCCCAACCGTCCCAACCGTCCCAACCGTCCCGCT 1260  
QY 1299 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCGCA 1358  
Db 1261 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCGCA 1320  
QY 1359 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCGCA 1418  
Db 1321 CTTCCACTTCAAGCCCTCGCGCGCATCCCGCGCGCATCCCGCGCGCGCGCGCGCGCA 1380  
QY 1419 CGCGGCGCAGCGCGCGCGCTTGAAGCACCAGTACCGGCTGCGCGCTGCGCGCGCGCG 1478  
Db 1381 CGCGGCGCAGCGCGCGCGCTTGAAGCACCAGTACCGGCTGCGCGCTGCGCGCGCGCG 1440  
QY 1479 GCGGCGCGCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 1538  
Db 1441 GCGGCGCGCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCT 1500  
QY 1539 GGTGGGCGAGTCCAGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1598  
Db 1501 GGTGGGCGAGTCCAGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1560  
QY 1599 CACGTGCGCGTGGGGGCAACCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1658  
Db 1561 CACGTGCGCGTGGGGGCAACCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1620  
QY 1659 GGGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCG 1718  
Db 1621 GGGCTGCAAGGGCTTTTCAAGAGAACAGTGCAGAAAAATGCAAAATATGTTTGCCTGCG 1680  
QY 1719 AATAAATACTCCAGTAGACAAGAGCTGCAAAACCGATGTCAAGTGTGCTGATTTCA 1778  
Db 1681 AATAAATACTCCAGTAGACAAGAGCTGCAAAACCGATGTCAAGTGTGCTGATTTCA 1740  
QY 1779 GAAGTGTCTCAGTGTGGAAATGCTAAAGAGTGTGCGTACAGATAGTCTGAAAGGGAG 1838  
Db 1741 GAAGTGTCTCAGTGTGGAAATGCTAAAGAGTGTGCGTACAGATAGTCTGAAAGGGAG 1800

QY 1839 GAGAGTGTCTGCTGCTTCCAAACCAAGAGCCCAATTACAACAGAGAACTTCTCAGCCCTC 1898  
Db 1801 GAGAGTGTCTGCTGCTTCCAAACCAAGAGCCCAATTACAACAGAGAACTTCTCAGCCCTC 1860  
QY 1899 TCCACCTTCTCTCTCAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAAC 1958  
Db 1861 TCCACCTTCTCTCTCAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAAC 1920  
QY 1959 ACCCAGAGATCTTGATTTATCCAGATAGTGCACCTGACAGGCTGCTCAGGCGACAGA 2018  
Db 1921 ACCCAGAGATCTTGATTTATCCAGATAGTGCACCTGACAGGCTGCTCAGGCGACAGA 1980  
QY 2019 TGCTGAGCATGTGCAACAATTTCTAACCTCTGACAGCTTCCATTTGATGATTTCCAGAAG 2078  
Db 1981 TGCTGAGCATGTGCAACAATTTCTAACCTCTGACAGCTTCCATTTGATGATTTCCAGAAG 2040  
QY 2079 CTGGGCGAAGAAATTTGGGATTTACTGATCTCCCAAGAAAGATCAGACATTTACTTTAT 2138  
Db 2041 CTGGGCGAAGAAATTTGGGATTTACTGATCTCCCAAGAAAGATCAGACATTTACTTTAT 2100  
QY 2139 TGAATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTTCCATCAGGTCACCACTGCTGA 2198  
Db 2101 TGAATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTTCCATCAGGTCACCACTGCTGA 2160  
QY 2199 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTCAGATCGACTTCACTGCTGCTGCTG 2258  
Db 2161 AGATAAGTTTGTGTTCTGCAATGGACTTGTCTCAGATCGACTTCACTGCTGCTGCTG 2220  
QY 2259 TGGGAGTGGCTCAGCTCTATTAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTGA 2318  
Db 2221 TGGGAGTGGCTCAGCTCTATTAAGACTTTTCTTAAATTTGAGAGCTTGAACCTTGA 2280  
QY 2319 TATCCAGCCTTAGCTGCTGCTCAGACTGAGCATGATCAGAGAAAGACATGGGTTAAA 2378  
Db 2281 TATCCAGCCTTAGCTGCTGCTCAGACTGAGCATGATCAGAGAAAGACATGGGTTAAA 2340  
QY 2379 AGAACCAAGAGAGTGAAGAGCTATGCAACAAGATCACAAGAGCTTTAAAGACCA 2438  
Db 2341 AGAACCAAGAGAGTGAAGAGCTATGCAACAAGATCACAAGAGCTTTAAAGACCA 2400  
QY 2439 GAGTAAAGGAGAGCTGAGAGCCACGAGTCCAGAGTCCAGAGCTGCTGCTGCTGCTGCTG 2498  
Db 2401 GAGTAAAGGAGAGCTGAGAGCCACGAGTCCAGAGTCCAGAGCTGCTGCTGCTGCTGCTG 2460  
QY 2499 GAGGAAAGATCTGACCCCTGCGCTCCAGCGCATCTTCTACCTGAGCTGGAAGACTTGGT 2558  
Db 2461 GAGGAAAGATCTGACCCCTGCGCTCCAGCGCATCTTCTACCTGAGCTGGAAGACTTGGT 2520  
QY 2559 GTCTCCACCTTCCATGATGACAGCTTCTCTGGAACCCCTACCTTCTTCTTCTTCTTCT 2618  
Db 2521 GTCTCCACCTTCCATGATGACAGCTTCTCTGGAACCCCTACCTTCTTCTTCTTCTTCT 2580  
QY 2619 AGTGGAGAGTGGCTGCTCTCTCTAGCAGCTGCTTCTAGCAGCTGCTTCTAGCAGCTGCT 2678  
Db 2581 AGTGGAGAGTGGCTGCTCTCTCTAGCAGCTGCTTCTAGCAGCTGCTTCTAGCAGCTGCT 2640  
QY 2679 TTTGGAAACCTTATCATTTCTGCTCTTCTTAAAGAGGAAAAAGAGCTCTCTGTAAGAGCA 2738  
Db 2641 TTTGGAAACCTTATCATTTCTGCTCTTCTTAAAGAGGAAAAAGAGCTCTCTGTAAGAGCA 2700  
QY 2739 AAGACTTTCTTTTCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2798  
Db 2701 AAGACTTTCTTTTCTTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2760  
QY 2799 TATTGCTGTTGGGTTGTTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2858  
Db 2761 TATTGCTGTTGGGTTGTTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2820  
QY 2859 AGTTTCATGAGGGTTTCTTAAAGAAATGCTAAAGAGCACTTTTGGACAAATGCTATCCAG 2918  
Db 2821 AGTTTCATGAGGGTTTCTTAAAGAAATGCTAAAGAGCACTTTTGGACAAATGCTATCCAG 2880  
QY 2919 CAGGAAAAAAGAGGATATATACTGTTTAAAACTCTTTCTGGGGATCCAAATATAGT 2978

```
Db 2881 CAGGAAAGGAGGATATATATCTGTTTAAACCTCTTTCGGGAATCCAATTATAGT 2940
Qy 2979 TGCCTTGTATTTAAAAAAGCAAGCCAGCGGTGTTTCGCCAGGATAGGATGTCTCTTA 3038
Db 2941 TGCCTTGTATTTAAAAAAGCAAGCCAGCGGTGTTTCGCCAGGATAGGATGTCTCTTA 3000
Qy 3039 AAGATTGCTCCCTTCAAAATATGCTTCTGTATCAAAAGTACGTATGTTGTCACAAAG 3098
Db 3001 AAGATTGCTCCCTTCAAAATATGCTTCTGTATCAAAAGTACGTATGTTGTCACAAAG 3060
Qy 3099 GCAGAAACTTCTCTTTAAATTCCTTCTCTTCTTTATTTTAACAATGGTGAAGATGGAGG 3158
Db 3061 GCAGAAACTTCTCTTTAAATTCCTTCTCTTCTTTTAACAATGGTGAAGATGGAGG 3120
Qy 3159 ATTACTCAAAATCAGACATGCCAAACAAATATGCTGTTTTCCTTCCATATAAAGTGC 3218
Db 3121 ATTACTCAAAATCAGACATGCCAAACAAATATGCTGTTTTCCTTCCATATAAAGTGC 3180
Qy 3219 AATTTTAAAGTCTGCTTACTAAGTCTGTTTATTAACCTCTCTTTATCTATATGG 3278
Db 3181 AATTTTAAAGTCTGCTTACTAAGTCTGTTTATTAACCTCTCTTTATCTATATGG 3240
Qy 3279 AATAAAGAGGAGGAGTATGTTAGCAATGACACGTTTAAATATCCCTAGCAGAGCTGT 3338
Db 3241 AATAAAGAGGAGGAGTATGTTAGCAATGACACGTTTAAATATCCCTAGCAGAGCTGT 3300
Qy 3339 GTTACCTTCCCTGTCGATCCCTCTGAGGTATGCCCATCAAGACTTTTAGGCCATTTC 3398
Db 3301 GTTACCTTCCCTGTCGATCCCTCTGAGGTATGCCCATCAAGACTTTTAGGCCATTTC 3360
Qy 3399 TTGATGGAACAGATCCCTGCGCTGATGTCAGCTATCCTGAAAGTGGATCAGATTATA 3458
Db 3361 TTGATGGAACAGATCCCTGCGCTGATGTCAGCTATCCTGAAAGTGGATCAGATTATA 3420
Qy 3459 AACTCGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCACAGAGTCCCTAAA 3518
Db 3421 AACTCGATTACATGTAACCTGTTTGGTGTGTTCTATCAACCCACAGAGTTCCTAAA 3480
Qy 3519 CTGCTTCAGTATAGTAACCTGATGTTATTCATTCAGAAAGCGCCATAGTCAAGTTGA 3578
Db 3481 CTGCTTCAGTATAGTAACCTGATGTTATTCATTCAGAAAGCGCCATAGTCAAGTTGA 3540
Qy 3579 GTATTTGATCCCTAGATAAGAAATGCAATCAGCAGGAACCTGTCATACAGGTTAAGCA 3638
Db 3541 GTATTTGATCCCTAGATAAGAAATGCAATCAGCAGGAACCTGTCATACAGGTTAAGCA 3600
Qy 3639 CCAGGACAAATAAGGATTTTATAGATATAATTTAAATTTTCTTA - TTGGTTAAGGAGA 3696
Db 3601 CCAGGACAAATAAGGATTTTATAGATATAATTTAAATTTTGGTAATTGGGTTAAGGAGA 3660
Qy 3697 -CAATTTTGGAGAGCAAGCAAA--TCTTTTAAAAAATAGTATGAATGTGAATACTAGA 3752
Db 3661 CCAATTTTGGAGAGCAAGCAAAATCTTTTAAAAAATAGTATGAATGTGAATACTAGA 3720
Qy 3753 AAAGATTTAAAAAATAGTATGATGTAGTACTAGGAAGGAT 3794
Db 3721 AAAGATTTAAGAAATAGTATGATGTAGTACTAGGAAGGAT 3762
```

## RESULT 8

```
ADF28866
ID ADF28866 standard; DNA; 4162 BP.
AC ADF28866;
XX
XX
DT 12-FEB-2004 (first entry)
XX
DE Human nor-1 nuclear receptor nucleic acid 1651190.
XX
KW Human; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;
KW cytosolic; haematopoietic cell; ds.
XX
```

```
OS Homo sapiens.
XX WO2003088812-A2.
XX 30-OCT-2003.
XX 15-APR-2003; 2003WO-US011804.
XX 17-APR-2002; 2002US-0373238P.
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX (UNIW ) UNIV WASHINGTON.
XX Mullican SE, Conneely OM, Milbrandt J;
XX WPI; 2003-854017/79.
XX GENBANK; 1651190.
XX Inhibiting proliferation of a hematopoietic cell, useful for treating or
XX preventing leukemia, comprises modulating the level or activity of nor-1
XX and/or nur77 nuclear receptor.
XX Disclosure; SEQ ID NO 1; 101pp; English.
XX The present sequence is that of human nor-1 nuclear receptor nucleic acid
XX 1651190. Nor-1 has been identified as a molecular target for therapeutic
XX intervention in the treatment of myeloid leukaemia. The invention
XX provides methods of inhibiting the proliferation of a haematopoietic stem
XX cell or a haematopoietic myeloid cell by modulating the level of a nor-1
XX and/or nur77 nuclear receptor. This involves increasing the level of the
XX receptor polypeptide or polynucleotide, e.g. by administration of a
XX vector comprising the polynucleotide. A claimed method of treating
XX leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
XX haematopoietic stem cell or myeloid cell. Also claimed are methods of
XX identifying an upregulator of expression of nor-1 and/or nur77, of
XX identifying a compound for the treatment of leukaemia, and of screening
XX for a compound for treatment of leukaemia, and a mouse model for
XX leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
XX acid sequences.
XX Sequence 4162 BP; 1017 A; 1224 C; 991 G; 930 T; 0 U; 0 Other;
XX
XX Query Match 88.4%; Score 3352; DB 10; Length 4162;
XX Best Local Similarity 90.8%; Pred. No. 0;
XX Matches 3779; Conservative 0; Mismatches 15; Indels 369; Gaps 5;
Qy 1 ATAAATGACGTGCCGAGAGAGCGAGAAACGCGACGCGGAGAGCGGAGTCTCTGCT 60
Db 1 ATAAATGACGTGCCGAGAGAGCGAGAAACGCGACGCGGAGAGCGGAGTCTCTGCT 60
Qy 61 CCGCCCCCACCCTCCAGCTCCCTGCTCTCTCTCTCTCCCTCCATACACAGCGGCTCA 120
Db 61 CCGCCCCCACCCTCCAGCTCCCTGCTCTCTCTCTCTCTCCCTCCATACACAGCGGCTCA 120
Qy 121 CACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 178
Db 121 CACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 179 ACTTGCCTCTCCCGCGCTCACACCCCTCTTTCCTGAGCCCTTTCGCGTGCAGCGCG 238
Db 181 ACTTGCCTCTCCCGCGCTCACACCCCTCTTTCCTGAGCCCTTTCGCGTGCAGCGCG 240
Qy 239 CGCGCAGCTGGAGCGCCCTCCCGGGTCACTTTGCAACGCTGACGCGTCCCGGAGTGC 298
Db 241 CGCGCAGCTGGAGCGCCCTCCCGGGTCACTTTGCAACGCTGACGCGTCCCGGAGTGC 300
Qy 299 CGTGAGGTGGAAACAGCGCGGCGATCTCTCCCTCTGTCTCAGCCCAAGCCAGACGCC 358
Db 301 CGTGAGGTGGAAACAGCGCGGCGATCTCTCCCTCTGTCTCAGCCCAAGCCAGACGCC 360
Qy 359 CGCGAACCCTCTCGGCTGTCTCTCCCATGATCGGATCGGATCGGATCGGATCGGATCGG 418
Db 361 CGCGAACCCTCTCGGCTGTCTCTCCCATGATCGGATCGGATCGGATCGGATCGGATCGG 420
```

QY	419	CTCAGCGCCTCCGGGAGCCGCTGGGCTTGATACCGCAGCCTTCCGGGACAGCAGCTGT	478
DB	421	CTCACGCGCTCCGGGAGCGCTGGCTTGATACCGCAGCCTTCCGGGACAGCAGCTGT	480
QY	479	GACTCCCCCAGTGTCAGATTTTCGGGACAGCTCTCTAGAAACTTCGTCTTAAAGACGGAAAC	538
DB	481	GACTCCCCCAGTGTCAGATTTTCGGGACAGCTCTCTAGAAACTTCGTCTTAAAGACGGAAAC	540
QY	539	CGCCACAGCACTCAAAAGCCACTTCGGGAGAGGGCAGCCGGCAAGCCCGGGGCCCTGAGC	598
DB	541	CGCCACAGCACTCAAAAGCCACTTCGGGAGAGGGCAGCCGGCAAGCCCGGGGCCCTGAGC	600
QY	599	CTGGACCCCTTAGCGGTGCCGGGAGCAGACTGCCGGCGCTTCGCCCTCCGCCGGAAGTCGCGCTC	658
DB	601	CTGGACCCCTTAGCGGTGCCGGGAGCAGACTGCCGGCGCTTCGCCCTCCGCCGGAAGTCGCGCTC	660
QY	659	CTCTCACTCTCAGGCTTCGGCTGGGAGAGACCCCGAGCCGCCACCATTCAGCGCGCGAAGAT	718
DB	661	CTCTCACTCTCAGGCTTCGGCTGGGAGAGACCCCGAGCCGCCACCATTCAGCGCGCGAAGAT	720
QY	719	ACCTCCAGATATGCCCTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	778
DB	721	ACCTCCAGATATGCCCTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT	780
QY	779	ATCGGGCGGAGACATACAGCTTCGGAATACACCAAGGAGATCATGAACCCCGACTACACCA	838
DB	781	ATCGGGCGGAGACATACAGCTTCGGAATACACCAAGGAGATCATGAACCCCGACTACACCA	840
QY	839	AGTGAACCATGGACCTTTGGCAGCACTGAGATCAAGCTACAGCCAACAAGTCCCTGCCCA	898
DB	841	AGTGAACCATGGACCTTTGGCAGCACTGAGATCAAGCTACAGCCAACAAGTCCCTGCCCA	900
QY	899	GCATCAGTACTTTCGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTCGGTGT	958
DB	901	GCATCAGTACTTTCGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCTCGGTGT	960
QY	959	ACCAATATGACGGCCCTTCATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACC	1018
DB	961	ACCAATATGACGGCCCTTCATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACC	1020
QY	1019	ATCACCAACCAACCAACCAACCAACCAATCACAGCAGCAGCATCAGCAGCCATCCA	1078
DB	1021	ATCACCAACCAACCAACCAACCAACCAATCACAGCAGCAGCATCAGCAGCCATCCA	1080
QY	1079	TTCTCTCAGCCTCCAGCCGGAGGAGAGTGCTGCCAGCACTCCATGTACTCAAGC	1138
DB	1081	TTCTCTCAGCCTCCAGCCGGAGGAGAGTGCTGCCAGCACTCCATGTACTCAAGC	1140
QY	1139	AGTCCCCACCGTCCACCCCAACAGCGCGCTTCCCGCCGAGGGGGGGCGTTATGGG	1198
DB	1141	AGTCCCCACCGTCCACCCCAACAGCGCGCTTCCCGCCGAGGGGGGGCGTTATGGG	1200
QY	1199	ACGAGGCACTGCCCTCGCGCCCGGGCTGCATCGCAACCCGCCCGCTGTGGACCCCGCGA	1258
DB	1201	ACGAGGCACTGCCCTCGCGCCCGGGCTGCATCGCAACCCGCCCGCTGTGGACCCCGCGA	1260
QY	1259	TGAAGGGGTCCCAACGGTGGCCGGCGCGCTTCCCGCTCTTCACCTTCAAGCCCTCGC	1318
DB	1261	TGAAGGGGTCCCAACGGTGGCCGGCGCGCTTCCCGCTCTTCACCTTCAAGCCCTCGC	1320
QY	1319	CGCCGATCCCCCGCCAGCCCGCGCGGCGGCCACCACTCGGTACAGCCCGACGG	1378
DB	1321	CGCCGATCCCCCGCCAGCCCGCGCGGCGGCCACCACTCGGTACAGCCCGACGG	1380
QY	1379	CCGCTGCGCGCTCAGCTGCGCTGGGAGCCGAGCCGCGCGGGCAGCCAGGCGCGCG	1438
DB	1381	CCGCTGCGCGCTCAGCTGCGCTGGGAGCCGAGCCGCGCGGGCAGCCAGGCGCGCG	1440
QY	1439	CGCTTGAGAGCCACCGTACGGGCTGCGCTGGCCAGAGGGGGCGCCCGCTGCGCTTCC	1498
DB	1441	CGCTTGAGAGCCACCGTACGGGCTGCGCTGGCCAGAGGGGGCGCCCGCTGCGCTTCC	1500



Db 2581 ATGGACTGTCCTGCATCGACTTCAGTCCTTCGTGGATTTGGGGAGTGGCTCGACTCTA 2640  
QY 2279 TTAAGACCTTTCTTAAATTTGCAGAGCTGAACCTTGTATATCAAGCTTTAGCCTGCC 2338  
Db 2641 TTAAGACCTTTCTTAAATTTGCAGAGCTGAACCTTGTATATCAAGCTTTAGCCTGCC 2700  
QY 2339 TGTGACACTGACATGATCACAAGAGACATGGGTTAAAGAACCAAGAGAGTGAAG 2398  
Db 2701 TGTGACACTGACATGATCACAAGAGACATGGGTTAAAGAACCAAGAGAGTGAAG 2760  
QY 2399 AGCTATGCAACCAAGATCAACAGCAGTTTAAAGACCAACAGAGTGAAGGACAGGCTCTGG 2458  
Db 2761 AGCTATGCAACCAAGATCAACAGCAGTTTAAAGACCAACAGAGTGAAGGACAGGCTCTGG 2820  
QY 2459 AGCCCAACGAGTCCAAGGTCCTGGGTCCTGGTAGAAGTGAAGGATCTGCACCTGG 2518  
Db 2821 AACCACCAAGTCCAAGGTCCTGGGTCCTGGTAGAAGTGAAGGATCTGCACCTGG 2880  
QY 2519 GCTCCAGCGCATCTTCTACCTGGAAGCTGGAAGATCTGGTGTCTCCACCTTCGATCATTTG 2578  
Db 2881 GCTCCAGCGCATCTTCTACCTGGAAGCTGGAAGATCTGGTGTCTCCACCTTCGATCATTTG 2940  
QY 2579 ACAAGCTCTTCTGGACACCTTCTTCTTAATCAGGAGCAGTGGAGCAGTGCCT 2638  
Db 2941 ACAAGCTCTTCTGGACACCTTCTTCTTAATCAGGAGCAGTGGAGCAGTGCCT 3000  
QY 2639 CCTCTCCTAGCACCTCTTCTACGAGCAAGAGGATAGGTTTGGAAACCTTATCATTTCC 2698  
Db 3001 CCTCTCCTAGCACCTCTTCTACGAGCAAGAGGATAGGTTTGGAAACCTTATCATTTCC 3060  
QY 2699 TGTCTCTTCTTAAGAGGAAAAGCAGCTCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2758  
Db 3061 TGTCTCTTCTTAAGAGGAAAAGCAGCTCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 3120  
QY 2759 GCTCTTTCTTCTTACACCTTAAGACAGAAATCTCAGAGTATCTGTGGGTTGCTGT 2818  
Db 3121 GCTCTTTCTTCTTACACCTTAAGACAGAAATCTCAGAGTATCTGTGGGTTGCTGT 3180  
QY 2819 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTTTCTAA 2878  
Db 3181 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTTTCTAA 3240  
QY 2879 GAAATTTGCTAAACAAGCACTTTTGACAAATGCTATCCAGCAGGAAAAAAGGATAATA 2938  
Db 3241 GAAATTTGCTAAACAAGCACTTTTGACAAATGCTATCCAGCAGGAAAAAAGGATAATA 3300  
QY 2939 TAACCTGTTTAAACCTTCTTCTGGGAAATCCAAATATAGTCTGTTGTTATTTAAACAA 2998  
Db 3301 TAACCTGTTTAAACCTTCTTCTGGGAAATCCAAATATAGTCTGTTGTTATTTAAACAA 3360  
QY 2999 GAAACAGCAAGGGTTGTTCCGACAGGTAGGATGTCTTAAAGATTTGGTCCCTTCAAAAT 3058  
Db 3361 GAAACAGCAAGGGTTGTTCCGACAGGTAGGATGTCTTAAAGATTTGGTCCCTTCAAAAT 3420  
QY 3059 ATGCTTCTGTTATCAAGAGTACGATATGGTGCACAAAGCAGAAACTTCTTTTAATT 3118  
Db 3421 ATGCTTCTGTTATCAAGAGTACGATATGGTGCACAAAGCAGAAACTTCTTTTAATT 3480  
QY 3119 TCCTTCTTCTTTATTTAAACAAATGTTCAAGATGGAGATTAACCTACAAATCAGACAT 3178  
Db 3481 TCCTTCTTCTTTATTTAAACAAATGTTCAAGATGGAGATTAACCTACAAATCAGACAT 3540  
QY 3179 GCGAAAAACAATATGGCTGTTTCTTCCATAAACAAGTGCATTTTAAAGTGTGCT 3238  
Db 3541 GCGAAAAACAATATGGCTGTTTCTTCCATAAACAAGTGCATTTTAAAGTGTGCT 3600  
QY 3239 TACTAAGTCTGTTTATTAACCTCTCTTTTATTTCTATATGGAATAAAGAGGAGGAGTCA 3298  
Db 3601 TACTAAGTCTGTTTATTAACCTCTCTTTTATTTCTATATGGAATAAAGAGGAGGAGTCA 3660  
QY 3299 TGTTAGCAAAATGACAGGTTAATATCCCTAGCAGAGCTGTTCACCTTCCCTGTGATC 3358

Db 3661 TGTTAGCAAAATGACAGGTTAATATCCCTAGCAGAGCTGTGTTCACCTTCCCTGTGATC 3720  
QY 3359 CTTTCTGAGGTATGCCCCATCAAGACTTTTATAGGCCATTTCTGATGGAACCAAGATCCCTG 3418  
Db 3721 CTTTCTGAGGTATGCCCCATCAAGACTTTTATAGGCCATTTCTGATGGAACCAAGATCCCTG 3780  
QY 3419 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTG 3478  
Db 3781 CCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTG 3840  
QY 3479 TTTTGGTGTGTTCTTATCAACCCCAAGAGTTCCTTAAACTTGTCTTCAAGTTATAGTAAC 3538  
Db 3841 TTTTGGTGTGTTCTTATCAACCCCAAGAGTTCCTTAAACTTGTCTTCAAGTTATAGTAAC 3900  
QY 3539 TCACTGGTATATTCATTCAGAAAGCCCATTAAGTCAAGTTCAGTTCAGTTCCTTCAAGTAAG 3598  
Db 3901 TCACTGGTATATTCATTCAGAAAGCCCATTAAGTCAAGTTCAGTTCAGTTCCTTCAAGTAAG 3960  
QY 3599 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACCAGGACAAATAGGATTTT 3658  
Db 3961 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGGTAAGCACCAGGACAAATAGGATTTT 4020  
QY 3659 TATAGATATAATTTAAATTTTCTTA - TTGGTTAAGGAGA - CAATTTTGGAGAGCAAGCA 3715  
Db 4021 TATAGATATAATTTAAATTTTCTTAATTTGGTAAATTTGGTAAAGGACCAATTTTGGAGAGCAAGCA 4080  
QY 3716 AA --- TCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAGATTTTAAAAAATAGTAT 3772  
Db 4081 AATCTTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAGATTTTAAAAAATAGTAT 4140  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
Db 4141 GAGTGTGAGTACTAGGAAGGAT 4162

## RESULT 9

ABU61892

ID ABL61892 standard; DNA; 4977 BP.

XX

AC ABL61892;

XX

DT 15-MAY-2002 (first entry)

XX

DE Colon adenocarcinoma related gene sequence SEQ ID NO:229.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;

KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

XX 30-MAY-2001; 2001WO-US010838.

PR 05-JUN-2000; 2000US-0209473P.

PR 05-JUN-2000; 2000US-0209531P.

PR 18-SEP-2000; 2000US-023133P.

PR 18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

PR 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

PR 22-SEP-2000; 2000US-0234567P.

PR 25-SEP-2000; 2000US-0234923P.

PR 25-SEP-2000; 2000US-0234924P.

PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

PR 25-SEP-2000; 2000US-0235280P.

PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 03-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
XX  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
PS Claim 1; SEQ ID NO 229; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
CC tumour  
XX  
SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;

Query Match 75.8%; Score 2875; DB 6; Length 4977;  
Best Local Similarity 98.1%; Pred. NO. 0;  
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

QY 733 CCCTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790  
DB 94 CCTGTGTCAGGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTGTCAGA 153  
QY 791 CATACAGCTCGGAATACACCGGAGATCATGAACCCGACCTACACCAAGCTGACCATGG 850  
DB 154 CATACAGCTCGGAATACACCGGAGATCATGAACCCGACCTACACCAAGCTGACCATGG 213

QY 851 ACCTTGGCAGCACTGAGATCAACGGCTACAGCCACACAGTCCCTGCCCCAGCATCAGTACCT 910  
DB 214 ACCTTGGCAGCACTGAGATCAACGGCTACAGCCACACAGTCCCTGCCCCAGCATCAGTACCT 273  
QY 911 TCCTGGAGGGCTACTCCAGGCAACTACGAACCTCAAGCCTTCTCTGGGTGTACCAATGACAGC 970  
DB 274 TTGTGGAGGGCTACTCCAGGCAACTACGAACCTCAAGCCTTCTCTGGGTGTACCAATGACAGC 333  
QY 971 GGCCCTTTGATCAAAAGTGGAGAGGGGGGGGGCCAGCTACCATCACCATCACCACCAACC 1030  
DB 334 GGCCCTTTGATCAAAAGTGGAGAGGGGGGGGGCCAGCTACCATCACCATCACCACCAACC 393  
QY 1031 ACCACCAACCACCAACCATCACCAGCAGAGCATACAGCAGCCATTCATTCCTCCAGCT 1090  
DB 394 ACCACCAACCACCAACCATCACCAGCAGAGCATACAGCAGCCATTCATTCCTCCAGCT 453  
QY 1091 CCAGCCGGAGGAGAGGTGCTGCCAGCAGCTCCATGTACTTCAAGCAGTCCCAACCGT 1150  
DB 454 CCAGCCGGAGGAGAGGTGCTGCCAGCAGCTCCATGTACTTCAAGCAGTCCCAACCGT 513  
QY 1151 CCAGCCCCACCAACCGCCGCTTCCCGCCGAGGGGGGGCTTATGGAGCAGGACACTGC 1210  
DB 514 CCAGCCCCACCAACCGCCGCTTCCCGCCGAGGGGGGGCTTATGGAGCAGGACACTGC 573  
QY 1211 CCTCGGCGCCGCTGATCGCAACCGCCGCTGCTGTGGAACCGCCGATGAAGGCGTCC 1270  
DB 574 CCTCGGCGCCGCTGATCGCAACCGCCGCTGCTGTGGAACCGCCGATGAAGGCGTCC 633  
QY 1271 CCAGGTGGCGGGCGGCTTCCCGCTTCTTCCACTTCAAGCCCTCGCGCGCGATCCCC 1330  
DB 634 CCAGGTGGCGGGCGGCTTCCCGCTTCTTCCACTTCAAGCCCTCGCGCGCGATCCCC 693  
QY 1331 CCGGCGCCAGCCCGCGCGGCCACCACTTCGGCTACGACCCGACGCGCCCTGCGCGCC 1390  
DB 694 CCGGCGCCAGCCCGCGCGGCCACCACTTCGGCTACGACCCGACGCGCGCCCTGCGCGCC 753  
QY 1391 TCAGCTTGCCTGGAGCCGACGCGCGCGGGAGCGCAGCGCGCGCTTGGAGACC 1450  
DB 754 TCAGCTTGCCTGGAGCCGACGCGCGCGGGAGCGCAGCGCGCGCTTGGAGGCC 813  
QY 1451 ACCGTAGCGGCTCGCTGGCCAAAGAGGGGGGGCGCTTCCCGCTTCCCGCTTCCCGCC 1510  
DB 814 ACCGTAGCGGCTCGCTGGCCAAAGAGGGGGGGCGCTTCCCGCTTCCCGCTTCCCGCC 873  
QY 1511 TCAGCCCTCCCTACCGGCTCCAGCTTGTGGGCGAGAGTCCAGCTGCGGTGCGCGCC 1570  
DB 874 TCAGCCCTCCCTACCGGCTCCAGCTTGTGGGCGAGAGTCCAGCTGCGGTGCGCGCC 933  
QY 1571 CCAGCAGAGCTCGCTGCTGGCAGGCGCAGTGTGCGGTGCGGGGACACCGCGCT 1630  
DB 934 CCAGCAGAGCTCGCTGCTGGCAGGCGCAGTGTGCGGTGCGGGGACACCGCGCT 993  
QY 1631 GCCAGCACTACGGCTGCGAACTTCGAGGGCTGCAAGGGCTTTTTCAGAGAACAGTGC 1690  
DB 994 GCCAGCACTACGGCTGCGAACTTCGAGGGCTGCAAGGGCTTTTTCAGAGAACAGTGC 1053  
QY 1691 AGAAAAATGCAAAAATATGTTTTCCTGGCAATAAAAACTGCCAGTAGACAAGAGACGTC 1750  
DB 1054 AGAAAAATGCAAAAATATGTTTTCCTGGCAATAAAAACTGCCAGTAGACAAGAGACGTC 1113  
QY 1751 GAAACCGATGTACGTACTGTGATTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1810  
DB 1114 GAAACCGATGTACGTACTGTGATTTCAAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1173  
QY 1811 TTGTCCGTACAGATAGTCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAAACCAAGAGCC 1870  
DB 1174 TTGTCCGTACAGATAGTCTGAAAGGAGGAGAGAGTGTGCTGCTTCCAAAACCAAGAGCC 1233  
QY 1871 CATTACAAAGAAACCTTCTCAGCCCTTCTCAGCTTCTCCTCAATTCGATGATGAATG 1930  
DB 1234 CATTACAAAGAAACCTTCTCAGCCCTTCTCAGCCCTTCTCCTCAATTCGATGATGAATG 1293  
QY 1931 CCCTGTGTCGAGCTTTTAAACAGACTCAACCCAGAGATCTTGAATTTATTCAGATACTGTC 1990

Db	1294	CTCTGTCCGAGCTTTAACAGACTCAACACCAGAGATCTTTGATTAATTCAGAGATACGTGTC	1353
Qy	1991	CCACTGCACAGGCTGCTGCAGGCACAGATGCTGAGCATGTGCACAAATTTCTACAACTCC	2050
Db	1354	CCACTGACCAGGCTGCTGCAGGCACAGATGCTGAGCATGTGCACAAATTTCTACAACTCC	1413
Qy	2051	TGACAGGCTCCATTTGATGTATPCACAGAGCTGGGCAGAAAAAGATTCGGGATTTACTGATC	2110
Db	1414	TGACAGGCTCCATTTGATGTATCCAGAAAGCTGGGCAGAAAAAGATTCGGGATTTACTGATC	1473
Qy	2111	TCGCCAAGAGAGATCAGACATTACTTATGNAATCAGCCTTTTGGAGCTGTTGTCTCTCA	2170
Db	1474	TCGCCAAGAGAGATCAGACATTACTTATGNAATCAGCCTTTTGGAGCTGTTGTGTCTCTCA	1533
Qy	2171	GACTTTTCCATCAGGTCAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTGTTGTC	2230
Db	1534	GACTTTTCCATCAGGTCAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTGTTGTC	1593
Qy	2231	TGCATTCGACTTCAGTGCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTTT	2290
Db	1594	TGCATTCGACTTCAGTGCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTTT	1653
Qy	2291	CTTTAAATTTGCAGAGCCTGNAACCTTGATATCCAAAGCCTTAGCCTGCTGCAGACTGA	2350
Db	1654	CTTTAAATTTGCAGAGCCTGNAACCTTGATATCCAAAGCCTTAGCCTGCTGCAGACTGA	1713
Qy	2351	GCATGATCACAGAAAGACATCGGTTTAAAGAACCAAAAGAGAGTTCGAAGAGCTATGCAACA	2410
Db	1714	GCATGATCACAGAAAGACATCGGTTTAAAGAACCAAAAGAGAGTTCGAAGAGCTATGCAACA	1773
Qy	2411	AGATCACAAAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTCGAGCCCAACCGAGT	2470
Db	1774	AGATCACAAAGCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTCGAGCCCAACCGAGT	1833
Qy	2471	CCAAGTCTCGGGTGCCTTGTTAGAACTGAGGAAAGATCTGGAACCTCGGGGCTCCAGGCGA	2530
Db	1834	CCAAGTCTCGGGTGCCTTGTTAGAACTGAGGAAAGATCTGGAACCTCGGGGCTCCAGGCGA	1893
Qy	2531	TCCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTGACAAGCTCTTCC	2590
Db	1894	TCCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTGACAAGCTCTTCC	1953
Qy	2591	TGGACACCCCTACCTTTCTAATCAGGAGCAGTGGAGCAGTAGGCTGCCTCTCTCTAGCA	2650
Db	1954	TGGACACCCCTACCTTTCTAATCAGGAGCAGTGGAGCAGTAGGCTGCCTCTCTCTAGCA	2013
Qy	2651	CCTGCTTGCTACGCAGCAAAAGGATAGGTTTGAAAACTTATCATTTCTGTCTTCTCTTA	2710
Db	2014	CCTGCTTGCTACGCAGCAAAAGGATAGGTTTGAAAACTTATCATTTCTGTCTTCTCTTA	2073
Qy	2711	AGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTTTCTGGCTCTTTTCTCTT	2770
Db	2074	AGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTTTCTGGCTCTTTTCTCTT	2133
Qy	2771	ACAACTTAAAGCCAGAAAACTTGACAGAGTATTTGTTTGGGGTGTGTTTTTATATTAGGC	2830
Db	2134	ACAACTTAAAGCCAGAAAACTTGACAGAGTATTTGTTTGGGGTGTGTTTTTATATTAGGC	2193
Qy	2831	ATTGGGGATGGGGTGGAGGGGTTATAGTTCATGAGGGTTTTCTAAGAAAATTGCTTAAC	2890
Db	2194	ATTGGGGATGGGGTGGAGGGGTTATAGTTCATGAGGGTTTTCTAAGAAAATTGCTTAAC	2253
Qy	2891	AAAGCACTTTTGGCAATGCTATCCACAGCAGGAAAAAAAAGGATAATATACTGTTTTAA	2950
Db	2254	AAAGCACTTTTGGCAATGCTATCCACAGCAGGAAAAAAAAGGATAATATACTGTTTTAA	2313
Qy	2951	AACCTTTCTGGGGAATCCAATTAAGTGTCTTTGTTATTTAAAAAACAAGAACAGCCAAAGG	3010
Db	2314	AACCTTTCTGGGGAATCCAATTAAGTGTCTTTGTTATTTAAAAAACAAGAACAGCCAAAGG	2373
Qy	3011	GTGTTTCGCCAGGGTAGGATGTGTTTAAAGATTGGTCCCTTGAAAAATATGCTTCTTGTA	3070

Db	2374	GTGTGTCGCCAGGGTAGGATGTGTCTTAAAGATTGGTCCCTCGAAATATATGCTTCCTGTA	2433
Qy	3071	TCAAAGGTACGATATGTGGTCAAAACAAGGCAGAAAACTTCCTTTTAATTTCTTCTTCCCTT	3130
Db	2434	TCAAAGGTACGATATGTGGTCAAAACAAGGCAGAAAACTTCCTTTTAATTTCTTCTTCCCTT	2493
Qy	3131	TATTTTAACAAAATGGTGAAGATGGAGGATTACCTACAAATCAGACATGCGCAAAACAATA	3190
Db	2494	TATTTTAACAAAATGGTGAAGATGGAGGATTACCTACAAATCAGACATGCGCAAAACAATA	2551
Qy	3191	ATGGCTGTTTGGCTTCCATAAACCAAGTGCAATTTTTTAAAGTGTCTTACTTAAGTCTTG	3250
Db	2552	ATGGCTGTTTGGCTTCCATAAACCAAGTGCAATTTTTTAAAGTGTCTTACTTAAGTCTTG	2608
Qy	3251	TTTATTAACTCTCTCTTTATTTCTATATATGGAATAAAAAAGGAGGCAGTCATGTTAGCAAAATG	3310
Db	2609	T---TATACCTCTTATCTATATATGCGGAATAAAAAAGGAGGCAGTCATGTTAGCAAAATG	2665
Qy	3311	ACAGCTTAATATCCCTAGCAGAGCGTGTGTTCACCTTCCCTGTGCATCCCTCTGAGGTA	3370
Db	2666	ACAGCTTAATATCCCTAGCAGAGCGTGTGTTCACCTTCCCTGTGCATCCCTCTGAGGTA	2725
Qy	3371	TGGCCCATCAAAGACTTTTAGGCCATCTTGATGGAACAGACATCCCTGCCTGACGTGTC	3430
Db	2726	TGGCCCATCAAAGACTTTTAGGCCATCTTGATGGAACAGACATCCCTGCCTGACGTGTC	2785
Qy	3431	AGCTATCCTGAAAGTGGATCAGATATATAACTGGATTACATGTAACTGTTTGGTTGTGT	3490
Db	2786	AGCTATCCTGAAAGTGGATCAGATATATAACTGGATTACATGTAACTGTTTGGTTGTGT	2845
Qy	3491	TCTATCAACCCCAACAGAGTTCCCTAACTTGCTTCAGTTATAGTAACCTGACCTGGTATAT	3550
Db	2846	TCTATCAACCCCAACAGAGTTCCCTAACTTGCTTCAGTTATAGTAACCTGACCTGGTATAT	2904
Qy	3551	TCATTCAGAAGCCCATTAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACAATGCAATC	3610
Db	2905	TCATTCAGAAGCCCATTAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACAATGCAATC	2959
Qy	3611	AGCAGGAACCTGGTCATACAGGGTAAGCACAGGGACAATAAGGATTTTTATAGATATAAT	3670
Db	2960	TCAGAGGACTGGTCATACAGGGTAAGCACAGGGACAATAAGGATTTTTATAGATATAAT	3019
Qy	3671	TTAATTTTGTATTGTTTAAAGGAGACAAATTTTGGAGACCAAGCAAA---TCTTTTAAA	3727
Db	3020	TTAATTTTGTATTGTTTAAAGGAGACAAATTTTGGAGACCAAGCAAAATCTCTTTTAAA	3079
Qy	3728	AAATAGTATGAATGTGAATCTAGAAAAGATTTAAAAATATAGTATGAGTGTGAGTACTAG	3787
Db	3080	AAATAGTATGAATGTGAATCTAGAAAAGATTTAAAAATATAGTATGAGTGTGAGTACTAG	3139
Qy	3788	GAAGGAT 3794	
Db	3140	GAAGGAT 3146	
RESULT 10			
ABK84229			
ID	ABK84229 standard; cDNA; 4977 BP.		
AC	ABK84229;		
XX			
XX	14-AUG-2002 (first entry)		
DT			
XX	Human cDNA differentially expressed in granulocytic cells #800.		
DE			
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
KW	viral infection; parasitic infection; protozoal infection;		
KW	fungal infection; sterile inflammatory disease; psoriasis;		
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
KW	cardiac reperfusion injury; renal reperfusion injury; ARDS;		
KW	adult respiratory distress syndrome; inflammatory bowel disease;		
KW	Crohn's disease; ulcerative colitis; periodontal disease;		
KW	granulocyte activation; chronic inflammation; allergy.		

```
XX Homo sapiens.
OS WO200228999-A2.
XX
XX 11-APR-2002.
XX
XX 03-OCT-2001; 2001WO-US030821.
XX
XX 03-OCT-2000; 2000US-0237189P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression of
XX genes associated with granulocyte activation, which serves as diagnostic
XX markers that is useful for monitoring disease states and drug toxicity.
XX
XX Claim 1; SEQ ID NO 800; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing the
XX expression level to an expression level in an unactivated GC, where
XX differential expression of Gs is indicative of GCA. Also included are
XX modulating (M2) GA by contacting GC with an agent that alters the
XX expression of at least one gene in Gs; (2) screening (M3) for an agent
XX capable of modulating GCA or an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease using the gene expression
XX profile; (3) detecting (M4) an inflammation (especially chronic) in a
XX tissue, an allergic response in a subject, exposure of a subject to a
XX pathogen or sterile inflammatory disease, by detecting the level of
XX expression in a sample of the tissue of gene(s) from Gs, where the level
XX of expression of the gene is indicative of inflammation; (4) treating
XX (M5) an inflammation (especially chronic) or in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease, by contacting a tissue having inflammation with an
XX agent that modulates the expression of gene(s) from Gs in the tissue. M1
XX is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
XX for screening an agent capable of modulating GCA preferably in an
XX inflammation in a tissue; M4 is useful for detecting an inflammation
XX (especially chronic) in a tissue, an allergic response in a subject,
XX exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
XX psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
XX cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
XX respiratory distress syndrome, inflammatory bowel disease, Crohn's
XX disease, ulcerative colitis, periodontal disease; also bacterial
XX infection, viral infection, parasitic infection, protozoal infection,
XX fungal infection and M5 is useful for treating one of the above
XX conditions. The present sequence represents a gene differentially
XX expressed in granulocytes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;
XX
XX Query Match 75.8%; Score 2875; DB 6; Length 4977;
XX Best Local Similarity 98.1%; Pred. No. 0;
XX Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;
XX
XX 733 CCCTCGCTCGAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGG--CGCAGA 790
XX |||
XX 94 CCGTGTCTGCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGTCGTC 153
XX |||
XX 791 CATACAGCTCGGAATACACCGGAGATCATGAACCCGACCTACACCAAGCTGACCATGG 850
XX |||
XX 154 CATACAGCTCGGAATACACCGGAGATCATGAACCCGACCTACACCAAGCTGACCATGG 213
XX |||
```

Db 1294 CTCTGTCCGAGCTTTAAAGACTCAACCCAGAGATCTTGATTTATTCAGATACTGTC 1353  
Qy 1991 CCATGACCAAGCTGCTCAGGCAAGATGCTGAGCATGTGCAACAATTTCTACAACTTCC 2050  
Db 1354 CCATGACCAAGCTGCTCAGGCAAGATGCTGAGCATGTGCAACAATTTCTACAACTTCC 1413  
Qy 2051 TGACAGCCTCCATTCGATGTATCCAGAGCTGGGAGGAGAAAGATTCGGGATTTACTGATC 2110  
Db 1414 TGACAGCCTCCATTCGATGTATCCAGAGCTGGGAGGAGAAAGATTCGGGATTTACTGATC 1473  
Qy 2111 TCCCAAGAGAGATCAGACATTTACTTATTTGAATCAGCCTTTTGGAGCTGTTTGCCTCA 2170  
Db 1474 TCCCAAGAGAGATCAGACATTTACTTATTTGAATCAGCCTTTTGGAGCTGTTTGCCTCA 1533  
Qy 2171 GACTTTCCATCAGGTCAGGTCGCTGAAGATAAGTTTGTGTTCTGCAATGSACTTGTCC 2230  
Db 1534 GACTTTCCATCAGGTCAGGTCGCTGAAGATAAGTTTGTGTTCTGCAATGSACTTGTCC 1593  
Qy 2231 TGATCGACTTCAGTGCCTTGTGATTTGGGAGTGGCTGACTCTATTAAAGACTTTT 2290  
Db 1594 TGATCGACTTCAGTGCCTTGTGATTTGGGAGTGGCTGACTCTATTAAAGACTTTT 1653  
Qy 2291 CCTTAAATTTGAGAGCTGAACCTTGTATATCCAAAGCCTTAGCCTGCTGTCAGACTGA 2350  
Db 1654 CCTTAAATTTGAGAGCTGAACCTTGTATATCCAAAGCCTTAGCCTGCTGTCAGACTGA 1713  
Qy 2351 GCATGATCACAGAAAGACATGGGTTTAAAGAACCAAAGAGAGTCCGAAGACTATGCAACA 2410  
Db 1714 GCATGATCACAGAAAGACATGGGTTTAAAGAACCAAAGAGAGTCCGAAGACTATGCAACA 1773  
Qy 2411 AGATCAAGAGCTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGCCACCGAGT 2470  
Db 1774 AGATCAAGAGCTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGCCACCGAGT 1833  
Qy 2471 CCAAGTCTCTGGTCCCTGGTAGAATCTGAGGAAGATCTGCACCTGGGCTCCAGCGCA 2530  
Db 1834 CCAAGTCTCTGGTCCCTGGTAGAATCTGAGGAAGATCTGCACCTGGGCTCCAGCGCA 1893  
Qy 2531 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCC 2590  
Db 1894 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCC 1953  
Qy 2591 TGGACACCTTACCTTTCTAAATCAGAGCAGTGGAGCAGTGCCTCTCTCCTAGCA 2650  
Db 1954 TGGACACCTTACCTTTCTAAATCAGAGCAGTGGAGCAGTGCCTCTCTCCTAGCA 2013  
Qy 2651 CCTGCTTCTAGCAGCAAGAGGATAGTTTGGAAACCTATCATTTCTGCTGCTTCTTCA 2710  
Db 2014 CCTGCTTCTAGCAGCAAGAGGATAGTTTGGAAACCTATCATTTCTGCTGCTTCTTCA 2073  
Qy 2711 AGAGGAAAGCAGCTCTCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTTTTCTCTT 2770  
Db 2074 AGAGGAAAGCAGCTCTCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTTTTCTCTT 2133  
Qy 2771 ACAACCTAAAGCCAGAAACCTTCAGAGATAATGCTGGGTTGTGTTTATATTTAGGC 2830  
Db 2134 ACAACCTAAAGCCAGAAACCTTCAGAGATAATGCTGGGTTGTGTTTATATTTAGGC 2193  
Qy 2831 ATTGGGGATGGGTTGGAGGGGTTATGTTTCATGAGGTTTCTAGAAATTCCTAAC 2890  
Db 2194 ATTGGGGATGGGTTGGAGGGGTTATGTTTCATGAGGTTTCTAGAAATTCCTAAC 2253  
Qy 2891 AAAGCATTCTTGGACAATGCTATCCAGCAGGAGGAGGATATATATCTGTTTAA 2950  
Db 2254 AAAGCATTCTTGGACAATGCTATCCAGCAGGAGGAGGATATATATCTGTTTAA 2313  
Qy 2951 AACTCTTTCTGGGAAATCAATATATGTTGCTTGTATTTAAAAACAAGAACAGCAAGG 3010  
Db 2314 AACTCTTTCTGGGAAATCAATATATGTTGCTTGTATTTAAAAACAAGAACAGCAAGG 2373  
Qy 3011 GTTGTTCGCCAGGTTAGGATGTTCTTAAAGATTTGGTCCCTTGAAAAATATGCTTCTGTA 3070

Db 2374 GTTGTTCGCCAGGTTAGGATGTTCTTAAAGATTTGGTCCCTTGAAAAATATGCTTCTGTA 2433  
Qy 3071 TCAAAAGTACGTATGTGTGCAAAACAAGCAGAAACTTCTCTTTTAAATTTCTTCTTCTT 3130  
Db 2434 TCAAAAGTACGTATGTGTGCAAAACAAGCAGAAACTTCTCTTTTAAATTTCTTCTTCTT 2493  
Qy 3131 TATTTTAAACAAATGGTGAAGATGAGGATTTACCTACAAATACAGACATGGCAAAACAATA 3190  
Db 2494 TATTTTAAACAAATGGTGAAGATGAGGATTTACCTACAAATACAGACATGGCAAAACAATA 2551  
Qy 3191 ATGGCTGTTTCTTCCATAAAACAAGTGCATTTTAAAGTGTCTCTTACTAAGTCTTGG 3250  
Db 2552 ATGGCTGTTTCTTCCATAAAACAAGTGCATTTTAAAGTGTCTCTTACTAAGTCTTGG 2608  
Qy 3251 TTTTAAATCTCTCTTCTTATCTATATGGAATAAAAGAGGAGCATGCTGTTAGCAAAATG 3310  
Db 2609 T---TATATCTCTTATCTATATGCGGAAATAAAAGAGGAGCATGCTGTTAGCAAAATG 2665  
Qy 3311 ACAGTTAAATATCCCTAGCAGAGGCTGTTCACCTCCCTGTCATCCCTTCTGAGGTA 3370  
Db 2666 ACAGTTAAATATCCCTAGCAGAGGCTGTTCACCTCCCTGTCATCCCTTCTGAGGTA 2725  
Qy 3371 TGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCTGACTGTCC 3430  
Db 2726 TGGCCCATCCAAAGACTTTTAGGCCATTTCTGATGGAACAGATCCCTGCTGACTGTCC 2785  
Qy 3431 AGCTATCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAATCTGTTTGGTGTGT 3490  
Db 2786 AGCTATCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAATCTGTTTGGTGTGT 2845  
Qy 3491 TCTATCAACCCCAACAGAGTTCCCTTAAACTTTCAGTTAGTAACTGACTGCTGATAT 3550  
Db 2846 TCTATCAACCCCAACAGAGTTCCCTTAAACTTTCAGTTAGTAACTGCTGATATA- 2904  
Qy 3551 TCATTTCAGAAAGCCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC 3610  
Db 2905 TCATTTCAGAAAGCCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC 2959  
Qy 3611 AGCAGGAATCGTTCATACAGGTAAGCACCAGGGAACAATAAGGATTTTATAGATAAT 3670  
Db 2960 TCAGAGGATCGTTCATACAGGTAAGCACCAGGGAACAATAAGGATTTTATAGATAAT 3019  
Qy 3671 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGAGCAAGAAA---TCTTTTAAA 3727  
Db 3020 TTAATTTTGTATTGTTTAAAGGACAAATTTTGGAGAGCAAGAAA---TCTTTTAAA 3079  
Qy 3728 AAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTATGAGTGTGACTACTAG 3787  
Db 3080 AAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTATGAGTGTGACTACTAG 3139  
Qy 3788 GAAGGAT 3794  
Db 3140 GAAGGAT 3146

## RESULT 11

ADL83261

ID ADL83261 standard; cDNA; 4977 BP.

XX ADL83261;

XX 17-JUN-2004 (first entry)

XX Human PRO84725 cDNA, SEQ ID 463.

XX Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;  
XX Antiallergic; Muscular; Neuroprotective; Nephrotropic; Antiinflammatory;  
KW Gene Therapy; PRO; B cell related disorder; cancer;  
KW immune-mediated inflammatory disease; human; gene; ss.

OS Homo sapiens.

XX WO2004024097-A2.

XX PD 25-MAR-2004.  
XX PF 15-SEP-2003; 2003WO-US029097.  
XX PR 16-SEP-2002; 2002US-0411392P.  
XX PA (GETH) GENENTECH INC.  
XX PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood WI;  
XX PI Wu TD;  
XX DR WPI; 2004-329389/30.  
XX DR P-PSDB; ADL83262.  
XX PT New PRO polypeptide, useful for diagnosing and treating a B cell related  
XX PT disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune  
XX PT mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.  
XX PS Claim 2; Fig 463; 695pp; English.  
XX CC The present invention relates to PRO proteins and their coding sequences.  
XX CC The PRO proteins are useful for diagnosing and treating a B cell related  
XX CC disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide  
XX CC antigen unresponsiveness, selective IgA deficiency, selective IgM  
XX CC deficiency, selective deficiency of IgG subclasses, immunodeficiency with  
XX CC hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's  
XX CC lymphoma, intermediate lymphoma, follicular lymphoma, type II  
XX CC hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic  
XX CC anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or  
XX CC ankylosing spondylitis. The PRO proteins are also useful for preparing a  
XX CC medicament for treating a condition that is responsive to the PRO  
XX CC protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO  
XX CC coding sequences are useful as hybridization probes in chromosome and  
XX CC gene mapping, in preparing PRO proteins, or in generating transgenic  
XX CC animals or knockout animals, which in turn are useful in the development  
XX CC and screening of therapeutically useful reagents.  
XX SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;  
Query Match 75.8%; Score 2875; DB 12; Length 4977;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;  
QY 733 CCCTGCGTCCAAAGCCCAATATAGCCCTTCCCTCCAGGTCCAGTTATGCGG--CGCAGA 790  
DB 94 CCTGTCGTCGAAGCCCAATATAGCCCTTCCCTCCAGGTTCAGTTATGCGGTGCTCAGA 153  
QY 791 CATACAGCTCGGATACACCGAGGAGATCATGAACCCCGACTACACCAAGCTGACCATGG 850  
DB 154 CATACAGCTCGGATACACCGAGGAGATCATGAACCCCGACTACACCAAGCTGACCATGG 213  
QY 851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGCTCCCTGCCAGCATCAGTACT 910  
DB 214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGCTCCCTGCCAGCATCAGTACT 273  
QY 911 TCGTGGAGGGCTACTCGAGCAACTACGAATCAAGCCTTCCCTGCGGTACCAATGCGAGC 970  
DB 274 TTGTGGAGGGCTACTCGAGCAACTACGAATCAAGCCTTCCCTGCGGTACCAATGCGAGC 333  
QY 971 GGCCTTGTATCAAGTGGAGGG 1030  
DB 334 GGCCTTGTATCAAGTGGAGGG 393  
QY 1031 ACCACACACACACACACATACACAGCAGCAGCATCAGCAGCATCCTTCTCCAGCCT 1090  
DB 394 ACCACACACACACACATACACAGCAGCAGCATCAGCAGCATCCTTCTCCAGCCT 453  
QY 1091 CCAGCCCGGAGGAGGAGGTGCTGCCAGCAGCCTCCATGCTTCAAGCAGTCCCAACCGT 1150  
DB 454 CCAGCCCGGAGGAGGAGGTGCTGCCAGCAGCCTCCATGCTTCAAGCAGTCCCAACCGT 513  
QY 1151 CCACCCCCACACAGCGGGGGCTTCCCCCGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1210

DB 514 CCACCCCCACACAGCGCGGCTTCCCCCGCAGCGGGGGGGTATTGGAGAGGACACTGC 573  
QY 1211 CCTCGGGCGCGGCTGCATCGCACCCCGCGCGCTGCTGGACCCCGCCCGCATGAAGCGGTCC 1270  
DB 574 CCTCGGGCGCGGCTGCATCGCACCCCGCGCGCTGCTGGACCCCGCCCGCATGAAGCGGTCC 633  
QY 1271 CCACGGTGGCGCGGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCGCGCGCATCCCC 1330  
DB 634 CCACGGTGGCGCGGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCGCGCGCATCCCC 693  
QY 1331 CGCGCCCGACCGCGCGGCGGCGCACCACTCGCTAGCAGCCGACCGCGCGCTCGCGCGC 1390  
DB 694 CGCGCCCGACCGCGCGGCGGCGCACCACTCGCTAGCAGCCGACCGCGCGCTCGCGCGC 753  
QY 1391 TCAGCCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCC 1450  
DB 754 TCAGCCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCC 813  
QY 1451 ACCGTAACGGGCTGCGCTGGCGCAAGAGGGCGCGCGCGCTTCCCGCTCTCGGCGC 1510  
DB 814 ACCGTAACGGGCTGCGCTGGCGCAAGAGGGCGCGCGCGCTTCCCGCTCTCGGCGC 873  
QY 1511 TCAGCCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCC 1570  
DB 874 TCAGCCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGAGCC 933  
QY 1571 CCACGAGGAGCTCGCTGCGCGAGGGCGCGCGCGCTTCCCGCGCGCAACCGCGCT 1630  
DB 934 CCACGAGGAGCTCGCTGCGCGAGGGCGCGCGCGCTTCCCGCGCGCAACCGCGCT 993  
QY 1631 GCCAGCACTACGGGCTGCGCAACCTGCGAGGGCTGCGAGGGCTTTCAGAGAACAGTGC 1690  
DB 994 GCCAGCACTACGGGCTGCGCAACCTGCGAGGGCTGCGAGGGCTTTCAGAGAACAGTGC 1053  
QY 1691 AGAAATGCAAAATATGTTGCTGCGCAATAAACTGCCAGTAGACAGAGAGCTC 1750  
DB 1054 AGAAATGCAAAATATGTTGCTGCGCAATAAACTGCCAGTAGACAGAGAGCTC 1113  
QY 1751 GAAACCGATGTCACTACTGTGATTCAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1810  
DB 1114 GAAACCGATGTCACTACTGTGATTCAGAGTGTCTCAGTGTGGAATGTTAAAGAG 1173  
QY 1811 TTGTCCGTACAGATAGTCTGAAGGGAGGAGAGTGTCTGCTTCCAAACCAAGAGCC 1870  
DB 1174 TTGTCCGTACAGATAGTCTGAAGGGAGGAGAGTGTCTGCTTCCAAACCAAGAGCC 1233  
QY 1871 CATTAACAGGAACCTTCTCAGCCCTCTCACCTTCTCCTCAATCTGCAATGATGATG 1930  
DB 1234 CATTAACAGGAACCTTCTCAGCCCTCTCACCTTCTCCTCAATCTGCAATGATGATG 1293  
QY 1931 CCCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGAATTTCCAGATCTGTC 1990  
DB 1294 CTCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGAATTTCCAGATCTGTC 1353  
QY 1991 CCATGACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAACAACTTCTACACCTCC 2050  
DB 1354 CCATGACAGGCTGCTGCGAGGACAGATGCTGAGCATGTGCAACAACTTCTACACCTCC 1413  
QY 2051 TGACAGCCTCCATTTGATGATATCCAGAGCTGGGCGAGAAAGATTCGGGATTTACTGATC 2110  
DB 1414 TGACAGCCTCCATTTGATGATATCCAGAGCTGGGCGAGAAAGATTCGGGATTTACTGATC 1473  
QY 2111 TCCCCAAGAAGATCAGACATTTATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCTCA 2170  
DB 1474 TCCCCAAGAAGATCAGACATTTATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCTCA 1533  
QY 2171 GACTTTCATCAGGTCAACACTGCTGAGATAGTTTGTCTTCTGCAATGACTTGTCC 2230  
DB 1534 GACTTTCATCAGGTCAACACTGCTGAGATAGTTTGTCTTCTGCAATGACTTGTCC 1593  
QY 2231 TGCACTCAGCTTTCAGTGCCTTCCGTGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTT 2290

Db 1594 TGCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTT 1653  
Qy 2291 CCTTAAATTTGAGAGCCCTGAACCTTGATATCCAAAGCCCTTAGCTGCCTGTGAGCACTGA 2350  
Db 1654 CCTTAAATTTGAGAGCCCTGAACCTTGATATCCAAAGCCCTTAGCTGCCTGTGAGCACTGA 1713  
Qy 2351 GCATCATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGCAGAGAGCTATGCAACA 2410  
Db 1714 GCATGATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGCAGAGAGCTATGCAACA 1773  
Qy 2411 AGATCACAGAGCTTTAAAGAACCAAGAGAGTGCAGAGAGTGCAGAGAGTGCAGAGT 2470  
Db 1774 AGATCACAGAGCTTTAAAGAACCAAGAGAGTGCAGAGAGTGCAGAGAGTGCAGAGT 1833  
Qy 2471 CCAAGGTCCTGGTGCCTCTGTAGAACTGAGGAAGATCTGACCCCTGGGCTTCAGAGCA 2530  
Db 1834 CCAAGGTCCTGGTGCCTCTGTAGAACTGAGGAAGATCTGACCCCTGGGCTTCAGAGCA 1893  
Qy 2531 TCTTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTCACAGAGCTCTTCC 2590  
Db 1894 TCTTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATTCACAGAGCTCTTCC 1953  
Qy 2591 TGGACACCTCTACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTCTAGCA 2650  
Db 1954 TGGACACCTCTACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTCTAGCA 2013  
Qy 2651 CCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTCTCTT 2710  
Db 2014 CCTGCTGCTAGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTCTCTT 2073  
Qy 2711 AGAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTGCTCTCTCTCTCTT 2770  
Db 2074 AGAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTCTGCTCTCTCTCTCTT 2133  
Qy 2771 ACACTCTTAAAGCAGCAAACTTGCAGAGTATGTTGGGGTGTGTTTATATATTAGGC 2830  
Db 2134 ACACTCTTAAAGCAGCAAACTTGCAGAGTATGTTGGGGTGTGTTTATATATTAGGC 2193  
Qy 2831 ATTGGGGATGGGTGGAGGGGTATAGTTTCATGAGGGTTTCTAAGAAATTCCTAAC 2890  
Db 2194 ATTGGGGATGGGTGGAGGGGTATAGTTTCATGAGGGTTTCTAAGAAATTCCTAAC 2253  
Qy 2891 AAAGCAGCTTTGGCAATGCTATCCAGCAGGMAAAAGGATATATACTGTTTAA 2950  
Db 2254 AAAGCAGCTTTGGCAATGCTATCCAGCAGGMAAAAGGATATATACTGTTTAA 2313  
Qy 2951 AACTCTTTCTGGGAATCCAAATTAAGTTGCTTTGTTTAAACCAAGAACAGCAAGG 3010  
Db 2314 AACTCTTTCTGGGAATCCAAATTAAGTTGCTTTGTTTAAACCAAGAACAGCAAGG 2373  
Qy 3011 GTTGTTCGCCAGGATGATGCTTAAAGATTGGTCCCTTGAAATATGCTTCTCTGTA 3070  
Db 2374 GTTGTTCGCCAGGATGATGCTTAAAGATTGGTCCCTTGAAATATGCTTCTCTGTA 2433  
Qy 3071 TCAAGGTACGTATGCTGCAACAGCAGCAAACTTCTTTTAAATTCCTTCTCTCTT 3130  
Db 2434 TCAAGGTACGTATGCTGCAACAGCAGCAAACTTCTTTTAAATTCCTTCTCTCTT 2493  
Qy 3131 TATTTTAAACAATGCTGAAAGATGAGGATTAACCTTACAAATCAGACATGGCAAAACAATA 3190  
Db 2494 TATTTTAAACAATGCTGAAAGATGAGGATTAACCTTACAAATCAGACATGGCAAAACAATA 2551  
Qy 3191 ATGGCTGTTTCTTCATATAAACAAGTCAATTTTAAAGTCTGCTTCTAAGTCTTG 3250  
Db 2552 ATGGCTGTTTCTTCATATAAACAAGTCAATTTTAAAGTCTGCTTCTAAGTCTTG 2608  
Qy 3251 TTTTAACTCTCTCTTATCTATATGGAATTAAGAGGAGCAGTCAATGTAAGATG 3310  
Db 2609 T---TATTAATCTCTTATCTATATGCGGAAATAAAGAGGAGCAGTCAATGTAAGATG 2665  
Qy 3311 ACAGCTTAATCTCTTACAGAGGCTGTGTTTCACTTCTCTGTCGATCCCTTCTGAGGTA 3370  
Db 2666 ACAGCTTAATCTCTTACAGAGGCTGTGTTTCACTTCTCTGTCGATCCCTTCTGAGGTA 2725

Qy 3371 TGGCCCATCCAGACTTTTAGGCCATTTCTGTATGGAACAGATCCCTGCCCTGACTGTCC 3430  
Db 2726 TGGCCCATCCAGACTTTTAGGCCATTTCTGTATGGAACAGATCCCTGCCCTGACTGTCC 2785  
Qy 3431 AGCTATCTCTGAAAGTGGATCAGATTATATAAATCTGGATTACATGTAATCTGTTTGGTTGTCT 3490  
Db 2786 AGCTATCTCTGAAAGTGGATCAGATTATATAAATCTGGATTACATGTAATCTGTTTGGTTGTCT 2845  
Qy 3491 TCTATCAACCCCAAGGAGTTCCTTAAACTTCTTCTGTTTATAGTAAGTCTGCTGATAT 3550  
Db 2846 TCTATCAACCCCAAGGAGTTCCTTAAACTTCTTCTGTTTATAGTAAGTCTGCTGATAT 2904  
Qy 3551 TCATTCAAGAGGAGTAAAGTCAAGTTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC 3610  
Db 2905 TCATTCAAGAGGAGTAAAGTCAAGTTCAGTTGAGTATTT--GATCCTAGATAAGAACATGCA-- 2959  
Qy 3611 AGCAGGAATCGTCTATCAGAGGTAAAGCAGGAGCAATTAAGGATTTTATAGATATAAT 3670  
Db 2960 TCAGAGGAGTCTGTCATCAGAGGTAAAGCAGGAGCAATTAAGGATTTTATAGATATAAT 3019  
Qy 3671 TTAATTTTGTATTTGTTTAAAGGAGCAATTTTGGAGAGCAAGCAAA--TCTTTTAAA 3727  
Db 3020 TTAATTTTGTATTTGTTTAAAGGAGCAATTTTGGAGAGCAAGCAAACTCTTCTTTTAAA 3079  
Qy 3728 AAATAGTATGAATGTGAATCTAGAAAAGATTTAAAAATAGTATGAGTGTGAGTACTAG 3787  
Db 3080 AAATAGTATGAATGTGAATCTAGAAAAGATTTTAAAGAAATAGTATGAGTGTGAGTACTAG 3139  
Qy 3788 GAAGGAT 3794  
Db 3140 GAAGGAT 3146

RESULT 12  
ADRI4062  
ID ADRI4062 standard; DNA; 4977 BP.  
AC ADRI4062;  
XX  
DT 21-OCT-2004 (first entry)  
XX  
DE Human NF-kappaB pathway-associated gene SeqID63.  
XX  
KW NF-kappaB pathway; antinflammatory; cytostatic; hepatotropic; virucide;  
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;  
KW immunosuppressive; vulnary; gene therapy; immune disorder;  
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;  
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
KW rheumatoid arthritis; host cell survival; evasion of immune response;  
KW atherosclerosis; cachexia; inflammatory bowel disease; colitis; asthma;  
KW autoimmune disorder; hyper immune activity;  
KW aberrant acute phase response; hypercongenital condition; birth defect;  
KW necrotic lesion; wound; organ transplant rejection;  
KW aberrant signal transduction; proliferating disorder; cancer;  
KW HIV propagation; gene; ds; human.  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO2004065577-A2.  
XX  
XX  
PD 05-AUG-2004.  
XX  
XX  
PF 13-JAN-2004; 2004WO-US000798.  
XX  
XX  
PR 14-JAN-2003; 2003US-0440068P.  
PR 12-MAY-2003; 2003US-0469757P.  
XX



PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Nadler SG, Neubauer MG, Feder JN, Carman J;  
XX  
XX WPI; 2004-562168/54.  
DR P-PSDB; ADRI4063.  
XX  
PT New isolated polynucleotides and polypeptides associated with NF-kappaB  
PT pathway, useful for diagnosing, treating, or preventing disorders or  
PT diseases associated with NF-kappaB pathway.  
XX  
PS Claim 1; SEQ ID NO 63; 237pp; English.  
XX  
CC This invention relates to the novel association of protein sequences (and  
CC the genes which encode them) to the NF-kappaB pathway. The invention may  
CC be useful for the production of compounds with an antiinflammatory,  
CC cyclostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
CC gastrointestinal-gen, antilasthmatic, antiarteriosclerotic,  
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
CC vulnary activity or for gene therapy. The proteins and nucleotides are  
CC useful for diagnosing, preventing, treating, or ameliorating conditions  
CC or diseases associated with the NF-kappaB pathway. The condition is an  
CC immune disorder, an inflammatory disorder, an inflammatory disorder  
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTIV-1,  
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
CC immune activity, disorders related to aberrant acute phase responses,  
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
CC organ transplant rejection, conditions related to organ transplant  
CC rejection, disorders related to aberrant signal transduction,  
CC proliferating disorders, cancers and HIV propagation in cells infected  
CC with other viruses. The present sequence is that of a human gene which is  
CC subject to the novel association with the NF-kappaB pathway of the  
CC invention. Note: This sequence does not appear in the specification but  
CC was obtained by the indexer from Genbank.  
XX  
SQ Sequence 4977 BP; 1394 A; 1147 C; 1042 G; 1394 T; 0 U; 0 Other;

Query Match 75.8%; Score 2875; DB 13; Length 4977;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

Qy 733 CCCTGGCTCCAGGCCCAATATAGCCCTTCCCTCCAGTTCAGTTATGCGG--CGCAGA 790  
Db |||  
Qy 94 CCTGTGTCCTCAAGGCCAATATAGCCCTTCCCTCCAGTTCAGTTATGCGGTCGTCAGA 153  
Db |||  
Qy 791 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG 850  
Db |||  
Qy 154 CATACAGCTCGGAATACACCGAGAGATCATGAACCCGAGCTACACCAAGCTGACCATGG 213  
Db |||  
Qy 851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACACACCTCCCTGCCAGCATCAGTACCT 910  
Db |||  
Qy 214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACACACCTCCCTGCCAGCATCAGTACCT 273  
Db |||  
Qy 911 TCGTGGAGGGCTACTCGAGCAACTACGAATCAGACCTTCTGGGTGACCAATGACG 970  
Db |||  
Qy 274 TTGTGGAGGGCTACTCGAGCAACTACGAATCAGACCTTCTGGGTGACCAATGACG 333  
Db |||  
Qy 971 GGCCCTTCATCAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1030  
Db |||  
Qy 334 GGCCCTTCATCAAGTGGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 393  
Qy 1031 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1090  
Db |||  
Qy 394 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 453  
Qy 1091 CCAGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1150  
Db |||

Db 454 CCAGCCCGAGGAGCAGAGGTGCTGCCAGACACTCCATGTACTTCAAGCAGTCCCCACCGT 513  
Qy 1151 CCACCCCCACACACCGCCGCGCTTCCCGCCCGCAGGCGGGGGCGTATATGGACAGGACACTGC 1210  
Db |||  
Db 514 CCACCCCCACACACCGCCGCGCTTCCCGCCCGCAGGCGGGGGCGTATATGGACAGGACACTGC 573  
Qy 1211 CTTGCGGCGCGCGGTGATCGACCCCGCGCGCTGCTTGGACCCCGCGGATGAAGGGGTCC 1270  
Db 574 CTTGCGGCGCGCGGTGATCGACCCCGCGCGCTGCTTGGACCCCGCGGATGAAGGGGTCC 633  
Qy 1271 CCAGGTTGGCGGCGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCCGCGCGGATCCCC 1330  
Db 634 CCAGGTTGGCGGCGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCCGCGCGGATCCCC 693  
Qy 1331 CCGGCGCCAGCCCGCGCGGCGGCCACCACTCGGCTACGACCCGACGCGCGCTGCGCGGC 1390  
Db 694 CCGGCGCCAGCCCGCGCGGCGGCCACCACTCGGCTACGACCCGACGCGCGCTGCGCGGC 753  
Qy 1391 TCAGCCTGCCGCTGGGAGCGCGACCGCGCGGCGAGCAGGCGCGCGCTTGAGAGCC 1450  
Db 754 TCAGCCTGCCGCTGGGAGCGCGACCGCGCGGCGAGCAGGCGCGCGCTTGAGAGCC 813  
Qy 1451 ACCGTTAGGCGTGCCTGCGCAAGAGGGCGGCCCGCTGGCGCTTCCCGCTCTCGGCC 1510  
Db 814 ACCGTTAGGCGTGCCTGCGCAAGAGGGCGGCCCGCTGGCGCTTCCCGCTCTCGGCC 873  
Qy 1511 TCAGCCTGCCGCTTACCGCGCTGCGAGCCTGCTGGGCGAGAGTCCAGCGCTGCGCGGC 1570  
Db 874 TCAGCCTGCCGCTTACCGCGCTGCGAGCCTGCTGGGCGAGAGTCCAGCGCTGCGCGGC 933  
Qy 1571 CCAGCAGAGCTGCTGCTGGCGAGGCGCGCTGTCGCTGCTGGGGGACAAACCGCGCT 1630  
Db 934 CCAGCAGAGCTGCTGCTGGCGAGGCGCGCTGTCGCTGCTGGGGGACAAACCGCGCT 993  
Qy 1631 GCCAGCCTAGCGGCTGCGACCTGCGAGGCGCTGCAAGGGCTTTTCAAGAGAACAGTGC 1690  
Db 994 GCCAGCCTAGCGGCTGCGACCTGCGAGGCGCTGCAAGGGCTTTTCAAGAGAACAGTGC 1053  
Qy 1691 AGAAAAATGCAAAATATGTTTGGCTGGCAATAAAAACTGCCAGTAGTACAGAGAGCTC 1750  
Db 1054 AGAAAAATGCAAAATATGTTTGGCTGGCAATAAAAACTGCCAGTAGTACAGAGAGCTC 1113  
Qy 1751 GAAACCGAGTCACTGCTGATTCAGAGTGTCTCAGTGTGGTGGTAAATGTTAAAGAG 1810  
Db 1114 GAAACCGAGTCACTGCTGATTCAGAGTGTCTCAGTGTGGTGGTAAATGTTAAAGAG 1173  
Qy 1811 TTGTCCGTACAGATAGTCTGAAAGGAGGAGGAGTCTGCTGCTTCCAAACCAAGAGCC 1870  
Db 1174 TTGTCCGTACAGATAGTCTGAAAGGAGGAGGAGTCTGCTGCTTCCAAACCAAGAGCC 1233  
Qy 1871 CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATCTGCAATGATG 1930  
Db 1234 CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATCTGCAATGATG 1293  
Qy 1931 CCCTTGTCCGAGCTTTTAAAGACTCAACCCAGAGATCTTGATTTATTCAGATCTGTC 1990  
Db 1294 CTCTTGTCCGAGCTTTTAAAGACTCAACCCAGAGATCTTGATTTATTCAGATCTGTC 1353  
Qy 1991 CCAGTCCAGGCTGCTGAGGAGCAGATGCTGAGCATGTGCAACAAATTTCTACACCTCC 2050  
Db 1354 CCAGTCCAGGCTGCTGAGGAGCAGATGCTGAGCATGTGCAACAAATTTCTACACCTCC 1413  
Qy 2051 TGACAGCCTCATTTGATGTATTCAGAGCTGGGAGAGAAAGATTCCGGGATTTTACTGATC 2110  
Db 1414 TGACAGCCTCATTTGATGTATTCAGAGCTGGGAGAGAAAGATTCCGGGATTTTACTGATC 1473  
Qy 2111 TCCCAAGAGATCAGACATTACTTGAATCAGCCTTTTGGAGCTGTTGCTCTCA 2170  
Db 1474 TCCCAAGAGATCAGACATTACTTGAATCAGCCTTTTGGAGCTGTTGCTCTCA 1533  
Qy 2171 GACTTTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGACTGTCC 2230  
Db 1534 GACTTTCCATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGACTGTCC 1593

```
QY 2231 TGCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTT 2290
Db 1594 TGCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTT 1653
QY 2291 CTTTAAATTTGCAGAGCCTGAACTTTGATATCCAAAGCCTTAGCTGCTGTGACAGCTGA 2350
Db 1654 CTTTAAATTTGCAGAGCCTGAACTTTGATATCCAAAGCCTTAGCTGCTGTGACAGCTGA 1713
QY 2351 GCATGATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGCAGAGCTATGCAACA 2410
Db 1714 GCATGATCACAGAAAGACATGGTTTAAAGAACCAAGAGAGTGCAGAGCTATGCAACA 1773
QY 2411 AGATCACAGACAGTTTAAAGAACCAAGAGAGTGCAGAGCTATGCAAGCTGAGT 2470
Db 1774 AGATCACAGACAGTTTAAAGAACCAAGAGAGTGCAGAGCTATGCAAGCTGAGT 1833
QY 2471 CCAAGTCTCTGGTGCCTTCGTGAGAACTGAGGAAGATCTGCAACCTGGGCTCCAGAGCA 2530
Db 1834 CCAAGTCTCTGGTGCCTTCGTGAGAACTGAGGAAGATCTGCAACCTGGGCTCCAGAGCA 1893
QY 2531 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTCGCAAGCTCTTCC 2590
Db 1894 TCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTCGCAAGCTCTTCC 1953
QY 2591 TGGACACCTTACCTTCTTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCA 2650
Db 1954 TGGACACCTTACCTTCTTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCA 2013
QY 2651 CTGCTCTCTACGCAAGAGGATAGTTTGGAAACCTTATCATTTCTGCTCTTCTTCTTA 2710
Db 2014 CTGCTCTCTACGCAAGAGGATAGTTTGGAAACCTTATCATTTCTGCTCTTCTTCTTA 2073
QY 2711 AGAGAAAGCAGCTCTGTGAGAAAGCAAGACTTTCTTTTCTGCTCTTCTTCTTCTT 2770
Db 2074 AGAGAAAGCAGCTCTGTGAGAAAGCAAGACTTTCTTTTCTGCTCTTCTTCTTCTT 2133
QY 2771 ACAACCTTAAGCCAGAAAACCTTGCAGAGTATGTGTGGGGTGTGTTTTATATTAGGC 2830
Db 2134 ACAACCTTAAGCCAGAAAACCTTGCAGAGTATGTGTGGGGTGTGTTTTATATTAGGC 2193
QY 2831 ATTGGGGATGGGGTGGAGGGGTTAGTTTCATGAGGGTTTCTAAGAAATTCGCTAAC 2890
Db 2194 ATTGGGGATGGGGTGGAGGGGTTAGTTTCATGAGGGTTTCTAAGAAATTCGCTAAC 2253
QY 2891 AAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAGGATATATACTGTGTTTAA 2950
Db 2254 AAAGCAGCTTTTGGCAATGCTATCCAGCAGGAAAGGATATATACTGTGTTTAA 2313
QY 2951 AACTCTTTCTGGGGAATCAATATATAGTTGCTTTGATTTTAAACAGAGCAGCCAGG 3010
Db 2314 AACTCTTTCTGGGGAATCAATATATAGTTGCTTTGATTTTAAACAGAGCAGCCAGG 2373
QY 3011 GTTGTTCGCCAGGTGAGTATGCTTAAAGATGCTTCCCTTGAAGATATGCTTCTGTA 3070
Db 2374 GTTGTTCGCCAGGTGAGTATGCTTAAAGATGCTTCCCTTGAAGATATGCTTCTGTA 2433
QY 3071 TCAAAGGTACGTATGTTGTCAAACAGAGCAGAAACTTCTTTTAAATTTCTTCTTCTT 3130
Db 2434 TCAAAGGTACGTATGTTGTCAAACAGAGCAGAAACTTCTTTTAAATTTCTTCTTCTT 2493
QY 3131 TATTTTAAACAATGGTGAAGATGAGGATTAACCTAACAATCAGACATGGCAACAATA 3190
Db 2494 TATTTTAAACAATGGTGAAGATGAGGATTAACCTAC - AATCAGACATGGCAACAATA - 2551
QY 3191 ATGGCTGTTTCTTCCATATAACAAGTCAATTTTTTAAAGTCTGCTTACTAGTCTTG 3250
Db 2552 ATGGCTGTTTCTTCC - -ATACACCTGCAATTTTTTAAAGTCTGCTTACTAGTCTTG 2608
QY 3251 TTTTAAACTCTCTCTTATTTCTATATGAAATAAAAGAGGAGCAGTATGTTAGCAATG 3310
Db 2609 T - -TATACTCTTATCTATATGCGGAATAAAAGAGGAGCAGTATGTTAGCAATG 2665
```

```
QY 3311 ACAGTTAATATCCCTAGCAGAGGCTGTGTTACCTTCCCTGCTCGATCCCTTCTGAGTA 3370
Db 2666 ACAGTTAATATCCCTAGCAGAGGCTGTGTTACCTTCCCTGCTCGATCCCTTCTGAGTA 2725
QY 3371 TGGCCCATCCCAAGACTTTTAGGCCATCTTGATGGAACAGATCCCTGCTGACTGTCC 3430
Db 2726 TGGCCCATCCCAAGACTTTTAGGCCATCTTGATGGAACAGATCCCTGCTGACTGTCC 2785
QY 3431 AGCTATCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 3490
Db 2786 AGCTATCTGAAAAGTGGATCAGATTATAAACTGGATTACATGTAACCTGTTTGGTGTGT 2845
QY 3491 TCTATCAACCCCAACAGAGTTCCTTAAACTTGCCTTACCTTACCTGATTAAGCTGCTATAT 3550
Db 2846 TCTATCAACCCCAACAGAGTTCCTTAAACTTGCCTTACCTTACCTGATTAAGCTGCTATAT - 2904
QY 3551 TCATTTCAGAAGCGCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAAGAACTGCAATC 3610
Db 2905 TCATTTCAGAAGCGCCATAAAGTCAGTTGAGTATTT - GATCCTAGATAAAGAACTGCAATC - - 2959
QY 3611 AGCAGGAACCTGGTCATACAGGGTAAGCACAGGGCAATAAGGATTTTATAGATATAAT 3670
Db 2960 TCAGAGGACTGGTCATACAGGGTAAGCACAGGGCAATAAGGATTTTATAGATATAAT 3019
QY 3671 TTAATTTTGTGTTTATGTTTAAAGGACAAATTTTGGAGGCAAGCAAA - - -TCTTTTAAA 3727
Db 3020 TTAATTTTGTGTTTATGTTTAAAGGACAAATTTTGGAGGCAAGCAAACTCTCTTTTAAA 3079
QY 3728 AAATAGTATGAATGTGAATACCTAGAAAAGATTTTAAAAAATAGTATGAGTGTGAGTACTAG 3787
Db 3080 AAATAGTATGAATGTGAATACCTAGAAAAGATTTTAAAGAAATAGTATGAGTGTGAGTACTAG 3139
QY 3788 GAAGGAT 3794
Db 3140 GAAGGAT 3146
```

## RESULT 13

ADI31602  
ID ADI31602 standard; cDNA; 2714 BP.

XX AC ADI31602;

XX DT 17-JUN-2004 (first entry)

XX DE Human cDNA #928.

XX KW Human; gene; ss; immunological response; immunopathological condition;  
Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;  
irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;  
acute monocytic leukaemia; antiinflammatory; antiaethmatic; antiulcer;  
osteopathic; antiarthritic; antirheumatic; cytostatic.

XX OS Homo sapiens.

XX FN US6607879-B1.

XX PD 19-AUG-2003.

XX PF 09-FEB-1998; 98US-00023655.

XX PR 09-FEB-1998; 98US-00023655.

XX PA (INCY-) INCYTE CORP.

XX PI Cocks BG, Stuart SG, Seilhamer JJ;

XX DR WPI; 2003-895307/82.

XX PT A composition comprising a plurality of cDNAs, useful for detecting  
altered expression of genes in an immunological response or for  
diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma  
or osteoarthritis.

XX Claim 1; SEQ ID NO 928; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs  
CC for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperosinophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence represents a human cDNA of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2714 BP; 596 A; 924 C; 680 G; 514 T; 0 U; 0 Other;

Query Match 70.9%; Score 2689.6; DB 11; Length 2714;  
Best Local Similarity 99.9%; Fred. No. 0;  
Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 132 TCACCTCGCACACAGACACAGCGCGCACACAGGCTCCGACACACACTTCGGCTCTCCC 191  
Db 19 TCACACAAACACACAGACACAGCGCGCACACAGGCTCCGACACACACTTCGGCTCTCCC 78

Qy 192 GCGGCTCACACCCCTCTTGGCTTGGAGCCCTTGGCGGTGACGCGCGCGCGCGAGCTGGA 251  
Db 79 GCGGCTCACACCCCTCTTGGCTTGGAGCCCTTGGCGGTGACGCGCGCGCGAGCTGGA 138

Qy 252 CGCCCTCCCGGGCTCAGTTTGCAACGCTGCGGTCGCGGAGTGGCGGTGGAGTGGGA 311  
Db 139 CGCCCTCCCGGGCTCAGTTTGCAACGCTGAGGTGCGCGGAGTGGCGGTGGAGTGGGA 198

Qy 312 ACAGCGCGGATCTCTCCCTCTGTCTACAGCCCAAGCCAGGAGCGCCCGCGGAACCTCTC 371  
Db 199 ACAGCGCGGATCTCTCCCTCTGTCTACAGCCCAAGCCAGGAGCGCCCGCGGAACCTCTC 258

Qy 372 GGCTGTGTCTCCATAGTGGGATCGAGATCCCGATCCCGACAGCGGTCTACCGGCTCG 431  
Db 259 GGCTGTGTCTCCATAGTGGGATCGAGATCCCGATCCCGACAGCGGTCTACCGGCTCG 318

Qy 432 GGAGCGGTGGGCTTGTACCGAGCCCTTCCGGGACAGAGCTGTGACCTCCCGCCCGAG 491  
Db 319 GGAGCGGTGGGCTTGTACCGAGCCCTTCCGGGACAGAGCTGTGACCTCCCGCCCGAG 378

Qy 492 TGCAGATTTCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAACCGCCACAGCACTC 551  
Db 379 TGCAGATTTCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAACCGCCACAGCACTC 438

Qy 552 AAAGCCCACTGGGAAGAGGAGCGCCCGCAAGCCCGGCGCTTGGAGCTTGGAGCCCTTACG 611  
Db 439 AAAGCCCACTGGGAAGAGGAGCGCCCGCAAGCCCGGCGCTTGGAGCTTGGAGCCCTTACG 498

Qy 612 GGTGCGGCGAGCACTGCGGGGCTTTCGCTTCGCGGAGAGCTCGGCTCTCTTACACTCTC 671  
Db 499 GGTGCGGCGAGCACTGCGGGGCTTTCGCTTCGCGGAGAGCTCGGCTCTCTTACACTCTC 558

Qy 672 AGCCTCGCTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGATACCTCTCAGATAT 731  
Db 559 AGCCTCGCTGGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGATACCTCTCAGATAT 618

Qy 732 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTTCCAGGTTCCAGTTATGGGGCGAGAC 791  
Db 619 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTTCCAGGTTCCAGTTATGGGGCGAGAC 678

Qy 792 ATACAGCTTCGGAATACACACAGGAGATCATGAACCCCGAGCTACACCAAGCTGACATGGA 851  
Db 679 ATACAGCTTCGGAATACACACAGGAGATCATGAACCCCGAGCTACACCAAGCTGACATGGA 738

Qy 852 CTTTGGCAGCACTGAGATACGGCTACAGCCACCACTCTCTCTGCCCAGCATGATACCTT 911  
Db 739 CTTTGGCAGCACTGAGATACGGCTACAGCCACCACTCTCTCTGCCCAGCATGATACCTT 798

Qy 912 CGTGGAGGGCTACTCGAGCAACTACGAACCTTTCAGGCTTCTGCTGTACCAAAATGACGG 971  
Db 799 CGTGGAGGGCTACTCGAGCAACTACGAACCTTTCAGGCTTCTGCTGTACCAAAATGACGG 858

Qy 972 GCCCTTGATCAAAAGTGGAGAGGGCGGGCGCCAGCTACCATCACCATCACCACCA 1031  
Db 859 GCCCTTGATCAAAAGTGGAGAGGGCGGGCGCCAGCTACCATCACCATCACCACCA 918

Qy 1032 CCACCAACACACACACATCACAGCAGCAGCATCAGAGCAGCATTCCTCTCAGGCTC 1091  
Db 919 CCACCAACACACACACATCACAGCAGCAGCATCAGAGCAGCATTCCTCTCAGGCTC 978

Qy 1092 CAGCCCGAGGAGCAGGTGCTGCCAGCACCTTCTTCAAGAGTCTCCACACGCTC 1151  
Db 979 CAGCCCGAGGAGCAGGTGCTGCCAGCACCTTCTTCAAGAGTCTCCACACGCTC 1038

Qy 1152 CACCCCAACACAGCGGCTTCCCGCGAGCGGGCGGCTTATGGGACGAGGACCTGCC 1211  
Db 1039 CACCCCAACACAGCGGCTTCCCGCGAGCGGGCGGCTTATGGGACGAGGACCTGCC 1098

Qy 1212 CTCGGCGCGGCTGCAATGCGCACCCGGCGGCTCTCTGACCCCGCATGAAGCGGCTCC 1271  
Db 1099 CTCGGCGCGGCTGCAATGCGCACCCGGCGGCTCTCTGACCCCGCATGAAGCGGCTCC 1158

Qy 1272 CACGGTGGCGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCATCCGCC 1331  
Db 1159 CACGGTGGCGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCATCCGCC 1218

Qy 1332 CGCGCCAGCGCGCGCGCGCCACACCTCTGCTCAGACCCGCGCGCTGCGGCT 1391  
Db 1219 CGCGCCAGCGCGCGCGCGCCACACCTCTGCTCAGACCCGCGCGCTGCGGCT 1278

Qy 1392 CAGCTTGGCTGGAGCGCGAGCGCGCGCGGAGCAGCAGCGCGCGCTTGAGAGCA 1451  
Db 1279 CAGCTTGGCTGGAGCGCGAGCGCGCGCGGAGCAGCAGCGCGCGCTTGAGAGCA 1338

Qy 1452 CCGGTAGCGGCTCGGCTGGCAAGAGCGCGCGCGCTTCCCGCTTCCCGCTCTCGGCT 1511  
Db 1339 CCGGTAGCGGCTCGGCTGGCAAGAGCGCGCGCGCTTCCCGCTTCCCGCTCTCGGCT 1398

Qy 1512 CACGCTCTCCCTTACCGCTCAGCCTCTCTGGGCGAGAGTCCAGCCTTCCGCTCGCGCT 1571  
Db 1399 CACGCTCTCCCTTACCGCTCAGCCTCTCTGGGCGAGAGTCCAGCCTTCCGCTCGCGCT 1458

Qy 1572 CAGCAGAGCTCTGCTTGGCGAGGAGCAGTGTGCGGTGCGGGGAGCAAGCGCGCTG 1631  
Db 1459 CAGCAGAGCTCTGCTTGGCGAGGAGCAGTGTGCGGTGCGGGGAGCAAGCGCGCTG 1518

Qy 1632 CCAGCCTTACCGGCTGCGAACCTCGAGGGGCTCAAGGGGCTTTTCAAGAGAAACAGTGCA 1691  
Db 1519 CCAGCCTTACCGGCTGCGAACCTCGAGGGGCTGCAAGGGCTTTTCAAGAGAAACAGTGCA 1578

Qy 1692 GAAATAATGCAAAATATGTTTCCCTGGGCAATAAATACTGCCAGTGTAGCAAGAGAGCTCG 1751  
Db 1579 GAAATAATGCAAAATATGTTTCCCTGGGCAATAAATACTGCCAGTGTAGCAAGAGAGCTCG 1638

Qy 1752 AAACCGAGTGTAGTGTCTGATTTTCAAGAGTGTCTAGTGTGGAATGTAAGAAGT 1811

```
Db 1639 AAACCGATGTAGTACTGTCTGATTTTCAAGATGTCTAGTGTGGAATGGTAAAGAGT 1698
Qy 1812 TGTCGGTACAGATAGTCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGCCC 1871
Db 1699 TGTCGGTACAGATAGTCTGAAAGGAGGAGAGGTGCTGCTGCTTCCAAACCAAGAGAGCCC 1758
Qy 1872 ATTACACAGAACTTCTGAGCCCTCTCCACCTTCTCTCCATCTGCATGATGAATGC 1931
Db 1759 ATTACACAGAACTTCTGAGCCCTCTCCACCTTCTCTCCATCTGCATGATGAATGC 1818
Qy 1932 CTTGTCCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATCCAGATACTGTCC 1991
Db 1819 CTTGTCCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATCCAGATACTGTCC 1878
Qy 1992 CACTGACAGCTGCTGCAGGCACAGATGCTGAGCATGTGCAACAATTTCTCAACCTCCT 2051
Db 1879 CACTGACAGCTGCTGCAGGCACAGATGCTGAGCATGTGCAACAATTTCTCAACCTCCT 1938
Qy 2052 GACAGCTCCATTTGATGTATTCAGAGCTGGCAGAGAAAGATTCGGGATTTACTGATCT 2111
Db 1939 GACAGCTCCATTTGATGTATTCAGAGCTGGCAGAGAAAGATTCGGGATTTACTGATCT 1998
Qy 2112 CCCAAAGAGATCAGACATTTACTTATTTGAATCAGCCCTTTTGGAGCTGTTTGTCTCAG 2171
Db 1999 CCCAAAGAGATCAGACATTTACTTATTTGAATCAGCCCTTTTGGAGCTGTTTGTCTCAG 2058
Qy 2172 ACTTTCCATCAGGTCAAACTGCTGGAAGATAAGTTTGTGTTCTGCAATGACATTTGCTCT 2231
Db 2059 ACTTTCCATCAGGTCAAACTGCTGGAAGATAAGTTTGTGTTCTGCAATGACATTTGCTCT 2118
Qy 2232 GCATCGACTTCAGTGCCTTCTGGAATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2291
Db 2119 GCATCGACTTCAGTGCCTTCTGGAATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2178
Qy 2292 CTTAAATTTGAGAGCTGACCTTGATATCCAGCTTTAGCCTGCTGCTGAGCAGCTGAG 2351
Db 2179 CTTAAATTTGAGAGCTGACCTTGATATCCAGCTTTAGCCTGCTGCTGAGCAGCTGAG 2238
Qy 2352 CATGATCACAGAAAGACATGGTGTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACA 2411
Db 2239 CATGATCACAGAAAGACATGGTGTAAAGAACCAAGAGAGTTCGAAGAGCTATGCAACA 2298
Qy 2412 GATCAAGAGCTTTAAAGACCCAGAGTAAGGACAGGCTCTGAGGCCACCGAGTC 2471
Db 2299 GATCAAGAGCTTTAAAGACCCAGAGTAAGGACAGGCTCTGAGGCCACCGAGTC 2358
Qy 2472 CAAGTCTCGGTGCTGCTGAGTACTGAGAGATCTGACCTTGGGCTTCCAGCGCAT 2531
Db 2359 CAAGTCTCGGTGCTGCTGAGTACTGAGAGATCTGACCTTGGGCTTCCAGCGCAT 2418
Qy 2532 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATGACAGCTCTTCTCT 2591
Db 2419 CTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATTCATGACAGCTCTTCTCT 2478
Qy 2592 GGACACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGTGCTCTCTCTCTAGCAC 2651
Db 2479 GGACACCTTACCTTTCTAATCAGGAGCAGTGGAGCAGTGTGCTCTCTCTCTAGCAC 2538
Qy 2652 CTGCTTGCTAGCAGCAAGGATAGTTTGGAAACCTATCATTTCTCTGCTCTCTCTTAA 2711
Db 2539 CTGCTTGCTAGCAGCAAGGATAGTTTGGAAACCTATCATTTCTCTGCTCTCTCTTAA 2598
Qy 2712 GAGGAAAGAGCTCTCTGAGAAAGCAAGACTTTCTTTTCTGCTCTCTTTTCTTAA 2771
Db 2599 GAGGAAAGAGCTCTCTGAGAAAGCAAGACTTTCTTTTCTGCTCTCTTTTCTTAA 2658
Qy 2772 CAACCTAAGCCAGAAAATCTGCAGAGTATTTGTTGGGGTGTGTTTATATTTA 2827
Db 2659 CAACCTAAGCCAGAAAATCTGCAGAGTATTTGTTGGGGTGTGTTTATATTTA 2714
```

RESULT 14

```
AAT16151
ID AAT16151 standard; cDNA to mRNA; 4400 BP.
XX
AC AAT16151;
DT 19-JUL-1996 (first entry)
XX
DE Apoptotic cerebral neuron nuclear receptor gene.
XX
KW Rat nuclear receptor gene; embryo; cerebral neuron; induction; apoptosis;
amplification; primer; PCR; Antisense; ribozyme; nerve; disease; ss.
XX
OS Rattus norvegicus.
XX
FH Key Location/Qualifiers
FT CDS 700..2586
FT /*tag= a
FT /product= "rat nuclear receptor"
```

JP08023980-A.

30-JAN-1996.

15-JUL-1994; 94JP-00164434.

15-JUL-1994; 94JP-00164434.

(TERU ) TERUMO CORP.

WPI; 1996-133421/14.

P-PSDB; AAR92057.

Nuclear receptor expressed in apoptosis of cerebral neuron - used to design anti-sense oligo:nucleotide(s) and ribozyme(s) for treatment of cerebral nerve disease etc.

Claim 1; Page 7-10; 11pp; Japanese.

This is the nucleotide sequence of the novel rat nuclear receptor gene designated NUR88. The gene was isolated from 16 day old rat embryo cerebral neurons induced into apoptosis by amplifying the sequence from cDNA using primers AAT28371-4. Antisense nucleic acids and ribozymes targeted to the gene can be used as agents for treating cerebral nerve diseases

SQ Sequence 4400 BP; 1055 A; 1239 C; 1034 G; 1072 T; 0 U; 0 Other;

Query Match	65.9%;	Score 2500;	DB 2;	Length 4400;
Best Local Similarity	83.4%;	Pred. No. 0;		
Matches 3131;	Conservative 0;	Mismatches 535;	Indels 90;	Gaps 22;
Qy 46	CGGAGTCTCTGCTTCCGCGCCGCCACCCCTCCAGCTCCTGCTCTCTCCGCTCCCAT	105		
Db 1	CCGAGTCTCTGCTTCCGCGCCGCCACCCCTCCAGGCTGCTCTCTCTCCGCTCCCAT	60		
Qy 106	ACACAGACGCGCTCACACCCGCTCCCTCATCTGCGACACACAGACGCGGCTCACAG	165		
Db 61	ACACAGACGCGCTCACACCCGCTCCCTCATCTGCGACACACAGACGCGGCTCACAC	120		
Qy 166	GCTCGGACACACACT--TCGCTCTCCGCGCGCTCACACCCCTTTCGCTTCCGCTT	223		
Db 121	GCTCGGACACACACTTCCACTCTCTCCGCGCGCTCACACCCCTTCTCTCGGCGCCCTC	180		
Qy 224	GCCGCT-----GCAGCGCGCGCGCGAGCTGAGCCCTCCCGGCTCACTTTTCAACG	278		
Db 181	GCCGCTGTCGCGCGCGCGCGAGCCGCGAGCGCCCTCAGGGCTCACTTTGCAACG	240		
Qy 279	CTGACGCTGCGCGAGTGGCGTGGAGGTGGGAAACAGCGGCGGATCTCTCCCTCTGTC	338		
Db 241	CTGACAGAGCGGCGAGTGGCGTGGAGGTGGGAAACGTCGCGACATCTTAGCCCTGTC	300		
Qy 339	ACAGCCCAAGCCAGGACGCCCGGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATC	398		

Db 301 GCAGCCGAGACTGGACG-CTGCGGAACCTCTCGCGCGCTCTCCCATGATGTTGGATC 359  
Qy 399 GCAGATCCCCACACAGCC---GTCACCGGCTCCGGAGCGCTGGGCTTTGTATACACGC 455  
Db 360 GCAGATCCCCAGCAGCGCGTGTCTACCGGCTCTGGAGCGCTGGTTTGTGACCGC 419  
Qy 456 AGCCCTTCGGGACAGCAGCTGTGACTCTCCCCCAGTGCAGATTTGGGACAGCTCTCTA 515  
Db 420 AGCCCTTCGGGACAGCAGCTGTGACTCTCCCCCAATCCAGATTTGGGGTCTCTCTA 479  
Qy 516 GAAACTGCTCTAAGACGGAAACCGCCACAGCACTCAAAGCCCACTCGGGAAGGGCAG 575  
Db 480 GAAACTGCTCTAAGACGGAAACCTCCACAGAACCCCAAAGCCCACTCGGGAAGGGCAG 539  
Qy 576 CCGGCAAGCCCGGCGCTGAGCTTGGACCTTAGCGGTGCGGCGAGCA-----CTGC 629  
Db 540 CCGCAAGCCCGGCGGTGAGCTTGNACCTCAACAGAGCGGCGCAGCAGCGGCGGC 599  
Qy 630 CGGCGCTTCGCTCGCGGAGCTCGGCTCTCTCTACACTCTCAGCCTCCGCTGGAGAGAC 689  
Db 600 GGCTGCTTCGCTATCCGAGCTCCCGGCTCTCTACACTCTCAGCCTCCGCTGGAGAGAC 659  
Qy 690 CCCAGCCCCACCAATCAGCGCGCAAGATACCTTCCAGATATGCCCTCGCTCCAGGCCA 749  
Db 660 CCCAGCCCCACCAATCAGCGCGCAAGATACCTTCCAGATATGCCCTCGCTGCAAGGCCA 719  
Qy 750 ATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACAC 809  
Db 720 ATATAGCCCTTCGCTCGGGGTCACTTATGCCAGCAGACTTATGGCTCGGAATACAC 779  
Qy 810 CACGAGATCATGAACCCCGACTACCAAGCTGACCATGAGACCTTGGCAGCACTGAGAT 869  
Db 780 CACAGAAATCATGAACCCCGACTATGCCAAGCTGACCATGAGACCTTGGTAGCAGGGAT 839  
Qy 870 CACGCTACAGCAGCAGCTCCCTGCCAGCATCAGTACTTCTGTGAGGCGCTACTCGAG 929  
Db 840 CATGCCACGCGCCAGCGTCCCTGCCAGCTTCAGTACTTCTATGGAGGCGTACCCCGAG 899  
Qy 930 CAATACGAACTCAAGCTTCTCGTGTGACCAATGCGAG-----CGGCGCTT 977  
Db 900 CAGCTCGAACTCAAGCCCTCTGCTGTACCAATGCGCGCTTCTGGGCGCTCGGCGCTT 959  
Qy 978 GATCAAAAGTGAGAGGGGGCGGCCAGTACCATCAACATCAACCAACCAACCA 1037  
Db 960 GATCAAGATGAAGAGGGTCGCGAGCATGGCTACCAACCAACCAACCAACCATCA 1019  
Qy 1038 CCACCAACCATCACAGCAGCAGCATCAGCAGCATCATTCCTCAGGCTCCAGGCC 1097  
Db 1020 TCATCAACCAACCAACCAAGCAG-----CAGCAGCGCTCCATTCCTCCTCGGCCC 1073  
Qy 1098 GGAGACGAGGTGTGCCAGCAGCTCCATGTACTTCAAGCAGTCCCAACCGCTCCAGCCC 1157  
Db 1074 CGAGGACGAGGTACTGCCAGCAGCTCATGTACTTCAAGCAGTCTCGCGCTCTAGCCC 1133  
Qy 1158 CACCAACCGGCTTCCCGCCGAGGGGGCGGTATGGACGAGGCACTGCCCTCGGC 1217  
Db 1134 GACCACTCAGGCTTCCCGCCGAGGGGGGGCGCTGTGGACGACGAGCTGCCCTCTGC 1193  
Qy 1218 GCGCGCTGATCGACCCGCGCGCTGTGAGCCCGCGCATGAGGGGTTCCCAAGGT 1277  
Db 1194 GCCTGGTGCATCGCTCGGAGCCGCTGCTGGACCCGCGCATGAGGGCAGTGCCTCCCAAT 1253  
Qy 1278 GCGCGCGCGGCTTCCGCTCTTCCACTTCAAGCCCTCGCGCGCATCCCCCGGCGC 1337  
Db 1254 GCGCGCTGCTCGGCGCTTCCGATCTTCTTCAAGCCCTCACCGCACACCTTCCGCGCC 1313  
Qy 1338 CAGCCCGCGCGCGCCACCACTTGGCTACGACCCCGAGCGCGCTCGCGCTCAGCCT 1397  
Db 1314 CAGCCCGCGCGCGCCACCACTTGGCTATGACCCCAACCGCGCAGCTGGGCTCAGTCT 1373  
Qy 1398 GCGCTGGAGCGCAGCGCGCGCGGAGCAGCGCGCGCTTGGAGCGCACCCGTA 1457  
Db 1374 ACCCTTGGAGCGCGCGCGCGCGGAGCAAGCTGCTGCGCTCGAGGGCCATCCGTA 1433

Qy 1458 CGGCTGCGCTGCGCCAAAGAGGGCGGCCGCTTCCGCTTCCGCTCTCGGCTTCCAGCC 1517  
Db 1434 CGGCTGCGCTGCGCCAAAGAGGACGGCCACGTTGACCTTCCCTCGCTGGGCTTCCAGC 1493  
Qy 1518 CTCCCTTACCGCTTCCAGCTGTGGGCGAGAGTCCAGAGCTGCGCTGCGCGCCAGCAG 1577  
Db 1494 GTCCCTTACCGCTTCCAGCTGTGGGAGAGAGCCAGCTTACCATCGCCACCAATAG 1553  
Qy 1578 GAGTCTGCTGCTGCGCGAGGGCAGCTGCGTGTGGGGAGACACCGCGCTTCCAGCA 1637  
Db 1554 GAGTCTCATATCCGCGGAGGGCAGCTGTGTGTGTGGGGGACAAATGCTGCTTCCAGCA 1613  
Qy 1638 CTACGGCTGCGAACTCTCGAGGGCTCAAGGGCTTTTCAAAGAGAACAGTGCAGAAAAA 1697  
Db 1614 CTACGGAGTCCGACACTCTCGAGGGCTCTTCAAGAGAACGGTGCAGAAAAA 1673  
Qy 1698 TGCAAAAATATGTTTTCCTGCGAAATAAAAACTGCCAGTAGACAAGAGACGCTGAAACCG 1757  
Db 1674 CGCAAAATATGTTTGTGCTTGGCAATAAAAACTGCCGTAGACAAGAGACGCTGAAATCG 1733  
Qy 1758 ATGTCAGTACTGTGATTTTTCAGAGTGTCTCAGTGTGGATGTAAAGAGTGTCTCG 1817  
Db 1734 ATGTCAGTACTGTGAGGTTTTCAGAAAGTGTCTCAGTGTGGGATGTGAGGAGTGTGCG 1793  
Qy 1818 TACAGATAGTCTCAAAAGGAGGAGGTCGCTCTGCTTCCAAACCAAGAGCCCATTACA 1877  
Db 1794 TACAGATAGTCTGAAAGGAGGAGGTCGCTCTGCTTCCAAACCAAGAGCCCATTACA 1853  
Qy 1878 ACAGGAACCTTTCAGGCTTCTCCACCTTCTCCTCAATCTGCAATGATGATGATGATGAT 1937  
Db 1854 ACAGGAGCCCTCGCAGGCTTCCCAACCTCTCCTCCGATCTGTATGATGAAACGCGCTTGT 1913  
Qy 1938 CCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTATTTCCAGAGTACTCTCCACTGA 1997  
Db 1914 CCGAGCTTTAACAGACGCAACGCCAGAGACTTGTATTTCTCCAGATACTGTCCACCGA 1973  
Qy 1998 CCAGGCTGCTCGAGGCAAGATGCTGAGCATGTGCAACAAATTTTACAACTCTCTGACAGC 2057  
Db 1974 CCAGGCCACTCGGGCACAGACGCTGAGCAGCTGACAGAGTCTTACAACTCTTGCAGGC 2033  
Qy 2058 CTCCATTTGATGATTCAGAGAGCTGGGAGAAAGATTCGGGATTTACTGATCTCCCAA 2117  
Db 2034 CTCCATCGAGTGTCCAGAAAGCTGGGAGAAAGATCCCGGATTCATGATCTCCCAA 2093  
Qy 2118 AGAAGATCAGACATTACTTATTTGAATCAGCCCTTTTGGAGCTGTTGTCTCAGACTTTC 2177  
Db 2094 AGAAGATCAGACGTTACTTATAGAAATCAGCCCTTTTGGAGCTGTTGTTCTTAGACTTTC 2153  
Qy 2178 CATCAGGTCAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTTGTCTTGCATTCG 2237  
Db 2154 TATCAGGTCAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGAGCTTGTCTTGCACCG 2213  
Qy 2238 ACTTCAGTGCCTTCGTGGATTTGGGAGTGGCTCGACTCTATTTAAAGACTTTTCTTAAA 2297  
Db 2214 ACTTCAGTGCCTTCGGGATTTGGGAGTGGCTCGACTTCCATTTAAAGACTTTTCTTAAA 2273  
Qy 2298 TTTTGCAGAGCTGAACCTTGATATCCAGGCTTAGCCTGTGCTGTCAGCACTGAGCATGAT 2357  
Db 2274 TTTGCAGAGCTGAACTTGTATCCAGGCTTAGCCTGTGCTGTCAGCACTGAGTATGAT 2333  
Qy 2358 CACAGAAAGACATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAAGATCAC 2417  
Db 2334 CACAGAGCATGGGTTTAAAGAACCAAGAGAGTGGAGGAGTATGCAACAAGATCAC 2393  
Qy 2418 AAGCAGTTTAAAGACACCAAGAGTAAGGAGCAGGCTCTGGAGGCCACCGAGTCCAAGGT 2477  
Db 2394 AAGCAGTTTAAAGACACCAAGAGTAAGGAGCAGGCTCTGGAGGCCCTCAGAGCCCAAGGT 2453  
Qy 2478 CCTGGGTCCTGGTGTAGAACTGAGAGAGATCTGACACCTTGGGCTTCCAGCGCATCTCTTA 2537  
Db 2454 CCTTCGCGCATGTTGGAACTGAGGAAAGATCTGCACCCAGGCGCTTCCAGCGTATCTCTTA 2513









**THIS PAGE BLANK (USPTO)**



Db ||||| 301 TGGAGTGGGAACAGCGCGCATCTCTCCCTCGTTCACAGCCCAAGCCAGGACGCGCG 360  
Qy ||||| 361 CGGAACCTCTCGGTGTCTCTCCATGAGTGCAGATCGAGCATCCCCACACGCGCT 420  
Db ||||| 361 CGGAACCTCTCGGTGTCTCTCCATGAGTGCAGATCGAGCATCCCCACACGCGCT 420  
Qy ||||| 421 CACGCCCTCCGGGAGCGCTGGCTGTGTACACCGCAGCCCTTCGGGACAGCAGCTGTGA 480  
Db ||||| 421 CACGCCCTCCGGGAGCGCTGGCTGTGTACACCGCAGCCCTTCGGGACAGCAGCTGTGA 480  
Qy ||||| 481 CTCCCCCGCAGTCAGATTTCCGGACAGCTCTCTAGAAACTCGCTCTTAAGACGGAACCG 540  
Db ||||| 481 CTCCCCCGCAGTCAGATTTCCGGACAGCTCTCTAGAAACTCGCTCTTAAGACGGAACCG 540  
Qy ||||| 541 CCACAGCACTCAAAAGCCACTCGCGAAGAGGCGAGCCCGCAAGCCCGGGCCCTGAGCCT 600  
Db ||||| 541 CCACAGCACTCAAAAGCCACTCGCGAAGAGGCGAGCCCGCAAGCCCGGGCCCTGAGCCT 600  
Qy ||||| 601 GGACCTTTAGCGGTGCCGGGCGAGCACTGCCGGCGCTTTGCCCTGCCCGGACGTCGCTCCT 660  
Db ||||| 601 GGACCTTTAGCGGTGCCGGGCGAGCACTGCCGGCGCTTTGCCCTGCCCGGACGTCGCTCCT 660  
Qy ||||| 661 CTTACACTCTCAGCTCGGTGGAGAGACCCCGAGCCCGCCACCATTCAGCGCGGAAGATAC 720  
Db ||||| 661 CTTACACTCTCAGCTCGGTGGAGAGACCCCGAGCCCGCCACCATTCAGCGCGGAAGATAC 720  
Qy ||||| 721 CTTCCAGATATGCCCTGCTGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTAT 780  
Db ||||| 721 CTTCCAGATATGCCCTGCTGCTCAAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTAT 780  
Qy ||||| 781 GCGGCGCAGACATACAGCTCGGAATACACCAAGGAGATCATGAACCCCGCATACACCAAG 840  
Db ||||| 781 GCGGCGCAGACATACAGCTCGGAATACACCAAGGAGATCATGAACCCCGCATACACCAAG 840  
Qy ||||| 841 CTGACATATGACCTTGGCAGCACTGAGATACGGCTACAGCCACCAAGCTTCTCGGTGTAC 900  
Db ||||| 841 CTGACATATGACCTTGGCAGCACTGAGATACGGCTACAGCCACCAAGCTTCTCGGTGTAC 900  
Qy ||||| 901 ATCAGTACTTCTGTTGAGGGCTACTCGAGCAACTACGAACCTCAAGCTTCTCGGTGTAC 960  
Db ||||| 901 ATCAGTACTTCTGTTGAGGGCTACTCGAGCAACTACGAACCTCAAGCTTCTCGGTGTAC 960  
Qy ||||| 961 CAAATGACGCGCCCTTTGATCAAAAGTGAGAGGGGGCGGCGCCAGCACTACCATCACCAT 1020  
Db ||||| 961 CAAATGACGCGCCCTTTGATCAAAAGTGAGAGGGGGCGGCGCCAGCACTACCATCACCAT 1020  
Qy ||||| 1021 CACCACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 1080  
Db ||||| 1021 CACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 1080  
Qy ||||| 1081 CCTCAGCTCCAGCCCGGAGAGAGGTGTGCGCCAGCACTCCATGTACTTCAAGCAG 1140  
Db ||||| 1081 CCTCAGCTCCAGCCCGGAGAGAGGTGTGCGCCAGCACTCCATGTACTTCAAGCAG 1140  
Qy ||||| 1141 TCCCCACGCTCCACCCCAACCAACCGCCCTTCCCGCGAGGGGGGGGCGTTATGGAC 1200  
Db ||||| 1141 TCCCCACGCTCCACCCCAACCAACCGCCCTTCCCGCGAGGGGGGGGCGTTATGGAC 1200  
Qy ||||| 1201 GAGGCACTGCCCTCGGCGCCCGGTGATCCACCGCGCCCGCTGCTGGAGCCCGCGATG 1260  
Db ||||| 1201 GAGGCACTGCCCTCGGCGCCCGGTGATCCACCGCGCCCGCTGCTGGAGCCCGCGATG 1260  
Qy ||||| 1261 AAGCGGTGCCACAGTGGCGCGCGGTCTTCCGCTCTTCCACTTCAAGCCCTCGCGG 1320  
Db ||||| 1261 AAGCGGTGCCACAGTGGCGCGCGGTCTTCCGCTCTTCCACTTCAAGCCCTCGCGG 1320  
Qy ||||| 1321 CCGCATCCCCCGCGCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
Db ||||| 1321 CCGCATCCCCCGCGCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
Qy ||||| 1381 GCTGCGCGCTCAGCTGCGCTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440

Db ||||| 1381 GCTGCGCGCTCAGCTGCGCTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
Qy ||||| 1441 CTTGAGAGCCACCGTACGGGCTGCCGCTGCGCAAGAGGGCGCGCGCGCTTCCCG 1500  
Db ||||| 1441 CTTGAGAGCCACCGTACGGGCTGCCGCTGCGCAAGAGGGCGCGCGCGCTTCCCG 1500  
Qy ||||| 1501 CTTCTCGGCTCAGCGCCCTCCCTTACCGCTGCGCTGCGCGAGAGTCCAGCGCTG 1560  
Db ||||| 1501 CTTCTCGGCTCAGCGCCCTCCCTTACCGCTGCGCTGCGCGAGAGTCCAGCGCTG 1560  
Qy ||||| 1561 CCGTGCAGCCCGCAGCAGGAGCTGCTGCTGCGGAGGCGACGCTGCTGCGGCGAC 1620  
Db ||||| 1561 CCGTGCAGCCCGCAGCAGGAGCTGCTGCTGCGGAGGCGACGCTGCTGCGGCGAC 1620  
Qy ||||| 1621 AACGCCCTCCAGCAGCACTACGGCGTGCAGAACCTGCGAGGGCTGCAAGGGCTTTTCAAG 1680  
Db ||||| 1621 AACGCCCTCCAGCAGCACTACGGCGTGCAGAACCTGCGAGGGCTGCAAGGGCTTTTCAAG 1680  
Qy ||||| 1681 AGAAGTGCAGAAAAATGCAAAATATGTTTGCCTGCAAAATAAAATCTGCCAGTAGAC 1740  
Db ||||| 1681 AGAAGTGCAGAAAAATGCAAAATATGTTTGCCTGCAAAATAAAATCTGCCAGTAGAC 1740  
Qy ||||| 1741 AAGAGCGTGAAGCCGATGTCAGTACTGCTGATTTTTCAGAGTCTCTCAGTGTGGAAATG 1800  
Db ||||| 1741 AAGAGCGTGAAGCCGATGTCAGTACTGCTGATTTTTCAGAGTCTCTCAGTGTGGAAATG 1800  
Qy ||||| 1801 GTAAAGAAAGTGTGCTGCTCAGATAGTCTGAAAGGAGGAGAGTCTGCTGCCCTTCAAA 1860  
Db ||||| 1801 GTAAAGAAAGTGTGCTGCTCAGATAGTCTGAAAGGAGGAGAGTCTGCTGCCCTTCAAA 1860  
Qy ||||| 1861 CCAAAGAGCCATTAACAGGAACCTTCTCAGCCCTCTCAGCCCTCTCCTTCTCCTTCAATCTGC 1920  
Db ||||| 1861 CCAAAGAGCCATTAACAGGAACCTTCTCAGCCCTCTCAGCCCTCTCCTTCTCCTTCAATCTGC 1920  
Qy ||||| 1921 ATGATGAAATCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGTATTTCC 1980  
Db ||||| 1921 ATGATGAAATCCCTTGTCCGAGCTTTAAACAGACTCAACACCCAGAGATCTTGTATTTCC 1980  
Qy ||||| 1981 AGATACCTGCTCCACTGACAGGCTGCTGAGGCAAGATGTCAGATGTCGAAATTC 2040  
Db ||||| 1981 AGATACCTGCTCCACTGACAGGCTGCTGAGGCAAGATGTCAGATGTCGAAATTC 2040  
Qy ||||| 2041 TACAACTCTGACAGCTCCTATGATGATTCAGAGCTGCGGAGAGAGATTTCCGGGA 2100  
Db ||||| 2041 TACAACTCTGACAGCTCCTATGATGATTCAGAGCTGCGGAGAGAGATTTCCGGGA 2100  
Qy ||||| 2101 TTTTACTGATCTCCCAAGAGATCAGACATTAATTAATTAATTAATTAATTAATTAATTAAT 2160  
Db ||||| 2101 TTTTACTGATCTCCCAAGAGATCAGACATTAATTAATTAATTAATTAATTAATTAATTAAT 2160  
Qy ||||| 2161 TTTTCTCTCAGACTTTCCATCAGGTCAAAACATCTGCTGAAAGATAGTTTGTGTTCTGCAAT 2220  
Db ||||| 2161 TTTTCTCTCAGACTTTCCATCAGGTCAAAACATCTGCTGAAAGATAGTTTGTGTTCTGCAAT 2220  
Qy ||||| 2221 GGACTTCTCTGATCAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Db ||||| 2221 GGACTTCTCTGATCAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280  
Qy ||||| 2281 AAGACTTTTCTTAAATTTTGCAGAGCTGAACCTTGAATTCAGAGCTTCCAGCTGCTGCTG 2340  
Db ||||| 2281 AAGACTTTTCTTAAATTTTGCAGAGCTGAACCTTGAATTCAGAGCTTCCAGCTGCTGCTG 2340  
Qy ||||| 2341 TCAGACTGAGCATGATCAGAGAGAGATGCTGGGTAAAGAAACCAAGAGAGTCCAGAG 2400  
Db ||||| 2341 TCAGACTGAGCATGATCAGAGAGAGATGCTGGGTAAAGAAACCAAGAGAGTCCAGAG 2400  
Qy ||||| 2401 CTATGAAAGATCAGAGAGATTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAG 2460  
Db ||||| 2401 CTATGAAAGATCAGAGAGATTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAG 2460  
Qy ||||| 2461 CCCACCGAGTCCAGGCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
Db ||||| 2461 CCCACCGAGTCCAGGCTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520

```
Qy 2521 CTCAGGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGAC 2580
Db 2521 CTCAGGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGAC 2580
Qy 2581 AAGCTCTTCTCGGACACCCCTACTTCTTAATCAGGAGCAGTGGAGCAGTGCCTCC 2640
Db 2581 AAGCTCTTCTCGGACACCCCTACTTCTTAATCAGGAGCAGTGGAGCAGTGCCTCC 2640
Qy 2641 TCTCCTAGCAGCTGCTGCTACGAGCAAGAGGATAGGTTTGGAAACCTATCATTTCTCG 2700
Db 2641 TCTCCTAGCAGCTGCTGCTACGAGCAAGAGGATAGGTTTGGAAACCTATCATTTCTCG 2700
Qy 2701 TCCTTCTTAGAGGAAAGCAGCTCTGTTAGAAAGCAAGCACTTCTTTTTTCTTGGC 2760
Db 2701 TCCTTCTTAGAGGAAAGCAGCTCTGTTAGAAAGCAAGCACTTCTTTTTTCTTGGC 2760
Qy 2761 TCTTTTCTTCAACACCTTAAGCCAGAAACCTTGACAGATATGTTGTTGGGGTGTGTTTT 2820
Db 2761 TCTTTTCTTCAACACCTTAAGCCAGAAACCTTGACAGATATGTTGTTGGGGTGTGTTTT 2820
Qy 2821 ATATTTAGGCATTTGGGGATCGGGTGGAGGGGTTATAGTTCATAGGGTTTTCTAAGA 2880
Db 2821 ATATTTAGGCATTTGGGGATCGGGTGGAGGGGTTATAGTTCATAGGGTTTTCTAAGA 2880
Qy 2881 AATTGCTAACAAAGCACTTTTGGCAATGCTATCCAGCAGCAAAAAAAGGATAATATA 2940
Db 2881 AATTGCTAACAAAGCACTTTTGGCAATGCTATCCAGCAGCAAAAAAAGGATAATATA 2940
Qy 2941 ACTGTTTTAAACCTTTCTCGGGATCCAAATATAGTTGCTTGTATTTAAACCAAGA 3000
Db 2941 ACTGTTTTAAACCTTTCTCGGGATCCAAATATAGTTGCTTGTATTTAAACCAAGA 3000
Qy 3001 ACAGCCAAAGGTTGTTCCGCAAGGTAGGATGTCTTAAAGATTGGTCCCTTGAATAAT 3060
Db 3001 ACAGCCAAAGGTTGTTCCGCAAGGTAGGATGTCTTAAAGATTGGTCCCTTGAATAAT 3060
Qy 3061 GCTTCTGTATCAAGGTACGTATGTGTGCAACAGGCAAGAACTTCTTTTAAATTC 3120
Db 3061 GCTTCTGTATCAAGGTACGTATGTGTGCAACAGGCAAGAACTTCTTTTAAATTC 3120
Qy 3121 CTCTCTCTTTATTTTAAACAAATGTTGAAAGATGAGGATTACCTTACAAATCAGACATGG 3180
Db 3121 CTCTCTCTTTATTTTAAACAAATGTTGAAAGATGAGGATTACCTTACAAATCAGACATGG 3180
Qy 3181 CAAACAAATAATGCTGTGCTTCCATAAACAAGTGCATTTTAAAGTGTGCTTTA 3240
Db 3181 CAAACAAATAATGCTGTGCTTCCATAAACAAGTGCATTTTAAAGTGTGCTTTA 3240
Qy 3241 CTAAGTCTTGTATTATTAATCTCTCTTTATTTCTATATGGAATATAAAGGAGGCAAGTCA 3300
Db 3241 CTAAGTCTTGTATTATTAATCTCTCTTTATTTCTATATGGAATATAAAGGAGGCAAGTCA 3300
Qy 3301 TTAGCAATGACAGTGAATATCTCTAGCAGAGCTGTGTTCACTTCCCTGTGATGCC 3360
Db 3301 TTAGCAATGACAGTGAATATCTCTAGCAGAGCTGTGTTCACTTCCCTGTGATGCC 3360
Qy 3361 TTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTGCC 3420
Db 3361 TTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATTTCTGATGAAACAGATCCCTGCC 3420
Qy 3421 CTGACTGTCCAGCTATCTCTGAAGTGGATCAGATTATAAATCGGATTAACATGTAATGTT 3480
Db 3421 CTGACTGTCCAGCTATCTCTGAAGTGGATCAGATTATAAATCGGATTAACATGTAATGTT 3480
Qy 3481 TTGGTGTGTTCTATCAACCCCAACAGAGTCCCTTAACTTGTCTCAGTTATAGTAAGT 3540
Db 3481 TTGGTGTGTTCTATCAACCCCAACAGAGTCCCTTAACTTGTCTCAGTTATAGTAAGT 3540
Qy 3541 ACTGGTATATTCAATTCAGAGGGCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAGAA 3600
Db 3541 ACTGGTATATTCAATTCAGAGGGCCATAAAGTCAGTTGAGTATTTGATCCCTAGATAAGAA 3600
```

```
Qy 3601 CATCAAAATCAGCAGGAACCTGGTCTATACAGGTTAAGCACCCAGGACAAATAGGATTTT 3660
Db 3601 CATCAAAATCAGCAGGAACCTGGTCTATACAGGTTAAGCACCCAGGACAAATAGGATTTT 3660
Qy 3661 TAGATATAAATTAATTTTATTTGTTATTTAGGAGACAATTTTGGAGAGCAAGCAAA--- 3717
Db 3661 TAGATATAAATTAATTTTATTTGTTATTTAGGAGACAATTTTGGAGAGCAAGCAAAATCT 3720
Qy 3718 TCTTTTAAAAAATAGTATGTAATGTGAATCTAGAAAAGATTTAAAAAATAGTATGAGTG 3777
Db 3721 TCTTTTAAAAAATAGTATGTAATGTGAATCTAGAAAAGATTTAAGAAATAGTATGAGTG 3780
Qy 3778 TGAGTACTAGGAAGGAT 3794
Db 3781 TGAGTACTAGGAAGGAT 3797

RESULT 2
US-09-949-016-905
; Sequence 905, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 905
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-905
```

```
Query Match 98.1%; Score 3722; DB 4; Length 3802;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

Qy 1 ATAATGACGTGCGGAGAGCAGCGAAGCGCAGCGCGGAGAGCGAGTCTCTGCT 60
Db 1 ATAATGACGTGCGGAGAGCAGCGAAGCGCAGCGCGGAGAGCGAGTCTCTGCT 60
Qy 61 CCGCGCCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCCATACACAGCGCGTCA 120
Db 61 CCGCGCCCGCCACCCCTCCAGCTCTGCTCTCTCCGCTCCCATACACAGCGCGTCA 120
Qy 121 CACCGCTCTCCCTACTCGCACACACAGCAAGCGGCGCACAGGCTCCG--CACAC 178
Db 121 CACCGCTCTCCCTACTCGAACACACAGCAAGCGGCGCACAGGCTCCGCGCACAC 180
Qy 179 ACTTCGCTCTCCCGGGGCTCACACCCCTCTTGCCCTGAGCCCTTGGCGGTGACGCGG 238
Db 181 ACTTCGCTCTCCCGGGGCTCACACCCCTCTTGCCCTGAGCCCTTGGCGGTGACGCGG 240
Qy 239 CGCGCAGCTGGACCGCCCTCCCGGGTCACTTTGCAACGCTGAGTGCCTGCGGAGTGGC 298
Db 241 CGCGCAGCTGGACCGCCCTCCCGGGTCACTTTGCAACGCTGAGTGCCTGCGGAGTGGC 300
Qy 299 CGTGGAGTGGGAACAGCGGGGCTCTCTCCCTCTGCTCAGCCCAAGCCAGCCAGCC 358
Db 301 CGTGGAGTGGGAACAGCGGGGCTCTCTCCCTCTGCTCAGCCCAAGCCAGCCAGCC 360
Qy 359 CGCGGAACCTCTCGGCTGTGCTCTCCCATGATCGGATCGGATCGGATCGGATCGGATCGG 418
Db 361 CGCGGAACCTCTCGGCTGTGCTCTCTCCCATGATCGGATCGGATCGGATCGGATCGG 420
```

QY 419 CTACCGCCTCCGGAGCCGCTGGGCTTGTACACCGAGCCCTTCCGGGACAGAGCTGT 478  
DB 421 CTACCGCCTCCGGAGCCGCTGGCTTGTACACCGAGCCCTTCCGGGACAGAGCTGT 480  
QY 479 GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 538  
DB 481 GACTCCCCCAGTGCAGATTTCCGGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 540  
QY 539 CGCCACAGCACTCAAGCCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTCAGC 598  
DB 541 CGCCACAGCACTCAAGCCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCGCTCAGC 600  
QY 599 GTGACCCCTTAGCGGTGCCGGGACGACTCGCGGCGCTTCCGCTTCGCGGACGCTCCGCTC 658  
DB 601 CTGACCCCTTAGCGGTGCCGGGACGACTCGCGGCGCTTCCGCTTCGCGGACGCTCCGCTC 660  
QY 659 CTCTACACTCTAGCCTTCGCTCGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGAT 718  
DB 661 CTCTACACTCTAGCCTTCGCTCGAGAGACCCCGAGCCCAACATTCAGCGCGCAAGAT 720  
QY 719 ACCCTCCAGATATGCCCTCGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 778  
DB 721 ACCCTCCAGATATGCCCTTCGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTT 780  
QY 779 ATGGGGGCGACATACAGCTCGGAATACACCGAGAGATCATGAACCCCGACTACACCA 838  
DB 781 ATGGGGGCGACATACAGCTCGGAATACACCGAGAGATCATGAACCCCGACTACACCA 840  
QY 839 AGCTGACATGGACCTTGGGAGCACTGAGATCAGGCTACAGCCACACGCTCCCTGCCA 898  
DB 841 AGCTGACATGGACCTTGGGAGCACTGAGATCAGGCTACAGCCACACGCTCCCTGCCA 900  
QY 899 GCATCAGTACCTTCGTGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCCCTGCGGT 958  
DB 901 GCATCAGTACCTTCGTGAGGGCTACTCGAGCAACTACGAACTCAAGCCTTCCCTGCGGT 960  
QY 959 ACCAAATGACGGCCCTTGTATCAAGTGGAGGAGGGCGGGCCCGAGCTACCATCACC 1018  
DB 961 ACCAAATGACGGCCCTTGTATCAAGTGGAGGAGGGCGGGCCCGAGCTACCATCACC 1020  
QY 1019 ATAC 1078  
DB 1021 ATAC 1080  
QY 1079 TTCTCCAGCCTTCAGCCCGGAGGACGAGGTGCTGCCAGACCTTCCATGTACTTCAAGC 1138  
DB 1081 TTCTCCAGCCTTCAGCCCGGAGGACGAGGTGCTGCCAGACCTTCCATGTACTTCAAGC 1140  
QY 1139 AGTCCCAACCGTCCACCCCAACCGCGGCTTCCCGCCCGAGCGGGGCGGTATGGG 1198  
DB 1141 AGTCCCAACCGTCCACCCCAACCGCGGCTTCCCGCCCGAGCGGGGCGGTATGGG 1200  
QY 1199 ACGAGGCACTGCTCCGGGCGCGGCTGCATCGCACCGCGCGCGCTGCTGACCCGCGA 1258  
DB 1201 ACGAGGCACTGCTCCGGGCGCGGCTGCATCGCACCGCGCGCGCTGCTGACCCGCGA 1260  
QY 1259 TGAAGGCGGTCCCAACGCTGCGCGCGGCTTCCCGCTTCCACCTTCAAGCCCTCGC 1318  
DB 1261 TGAAGGCGGTCCCAACGCTGCGCGCGGCTTCCCGCTTCTTCCACTTCAAGCCCTCGC 1320  
QY 1319 CGCGCATCCCGCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1378  
DB 1321 CGCGCATCCCGCGCGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380  
QY 1379 CGCTGCGCGCTCAGCCTGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1438  
DB 1381 CGCTGCGCGCTCAGCCTGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
QY 1439 CGCTTGAGAGCACCCGCTACGCGCTGCGCTGGCCAGAGGCGCGCGCGCTGCGCTTCC 1498  
DB 1441 CGCTTGAGGCGCACCCGTAAGGCTGCGCTGGCCAGAGGCGCGCGCGCGCGCGCTTCC 1500

QY 1499 CGCCTCTCGGCTCACGCGCTCCCTACCGGCTCCAGCTGCTGGGGAGAGTCCCGAGCC 1558  
DB 1501 CGCCTCTCGGCTCACGCGCTCCCTACCGGCTCCAGCTGCTGGGGAGAGTCCCGAGCC 1560  
QY 1559 TGGCGTCCGCGCCAGCAGAGCTCGTCTGCGGAGGCGACGCTGCGCTGTCGCGGG 1618  
DB 1561 TGGCGTCCGCGCCAGCAGAGCTCGTCTGCGGAGGCGACGCTGCGCTGTCGCGGG 1620  
QY 1619 ACAACGCGGCTCCAGCACTACGGGCTGCGAACTCTGCGAGGCTGCAAGGCTTTTCA 1678  
DB 1621 ACAACGCGGCTCCAGCACTACGGGCTGCGAACTCTGCGAGGCTGCAAGGCTTTTCA 1680  
QY 1679 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTCCCTGGCAATAAATAATGCGCCAGTAG 1738  
DB 1681 AGAGAACAGTGCAGAAAAATGCAAAATATGTTTCCCTGGCAATAAATAATGCGCCAGTAG 1740  
QY 1739 ACAAGAGAGCTGCAAAACCGGATGTCAGTCTCGATTTCAAGAGTGTCTCAGTCTTCGAA 1798  
DB 1741 ACAAGAGAGCTGCAAAACCGGATGTCAGTCTCGATTTCAAGAGTGTCTCAGTCTTCGAA 1800  
QY 1799 TGGTAAAGAGAGTTGTCGCTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCCA 1858  
DB 1801 TGGTAAAGAGAGTTGTCGCTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCCA 1860  
QY 1859 AACCAAGAGCCATTAACAACAGAACCTTCTCAGCCCTCTCAGCTTCTCCTCCAATCT 1918  
DB 1861 AACCAAGAGCCATTAACAACAGAACCTTCTCAGCCCTCTCAGCTTCTCCTCCAATCT 1920  
QY 1919 GCATGATGAATGCCCTTGTCCGAGCTTTAACAAGACTCAACACCCAGAGATCTTGATTT 1978  
DB 1921 GCATGATGAATGCCCTTGTCCGAGCTTTAACAAGACTCAACACCCAGAGATCTTGATTT 1980  
QY 1979 CCAGATACTGTCCCACTGACAGGCTGCTCAGGSCACAGATGTGAGCATGTGCAACAAT 2038  
DB 1981 CCAGATACTGTCCCACTGACAGGCTGCTCAGGSCACAGATGTGAGCATGTGCAACAAT 2040  
QY 2039 TCTACAACCTCTGACAGCCCTCAATGATATCCAGAGCTGGGCAAGAGATTCGG 2098  
DB 2041 TCTACAACCTCTGACAGCCCTCAATGATATCCAGAGCTGGGCAAGAGATTCGG 2100  
QY 2099 GATTTACTGATCTCCCAAGAGATCAGACATTTACTTTAATCAGACCTTTTGGAGC 2158  
DB 2101 GATTTACTGATCTCCCAAGAGATCAGACATTTACTTTAATCAGACCTTTTGGAGC 2160  
QY 2159 TGTGTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTGTGTCTGCA 2218  
DB 2161 TGTGTCTCAGACTTTCCATCAGGTCAACACTGCTGAAGATAAGTTGTGTCTGCA 2220  
QY 2219 ATGGACTGTCTGCATCGACTTCAGTCCCTTGTGATTTGGGAGTGGCTCGACTCTA 2278  
DB 2221 ATGGACTGTCTGCATCGACTTCAGTCCCTTGTGATTTGGGAGTGGCTCGACTCTA 2280  
QY 2279 TTAAGACTTTTCTTAAATTTGCAGAGCTTGAACCTTGTATATCAAGCTTASCTTCC 2338  
DB 2281 TTAAGACTTTTCTTAAATTTGCAGAGCTTGAACCTTGTATATCAAGCTTASCTTCC 2340  
QY 2339 TGTCAAGCACTGAGCATCATCAGAAAGACATGGGTTAAAGAAACCAAGAGTCTGAAG 2398  
DB 2341 TGTCAAGCACTGAGCATCATCAGAAAGACATGGGTTAAAGAAACCAAGAGTCTGAAG 2400  
QY 2399 AGCTATCGAACAGATCACAAGCAGTTTAAAGACCAACAGAGTAAGGAGCAGCTCTGG 2458  
DB 2401 AGCTATCGAACAGATCACAAGCAGTTTAAAGACCAACAGAGTAAGGAGCAGCTCTGG 2460  
QY 2459 AGCCCAACGAGTCCAAAGGCTCTGGGTCCTTGGTAGAACTGAGGAAGATCTGCAACCTTGG 2518  
DB 2461 AACCAACGAGTCCAAAGTCTCTGGTTGCCCTTGTGTAAGTGAAGATCTGCAACCTTGG 2520  
QY 2519 GCCTTCAGAGCATCTTCTAAGCTGAGAGATTTGGTGTCTCCACCTTCCATCATTTG 2578  
DB 2521 GCCTTCAGAGCATCTTCTAAGCTGAGAGATTTGGTGTCTCCACCTTCCATCATTTG 2580  
QY 2579 ACAAGCTCTTCTCGACACACCTTCTTCTAATCAGGAGCAGTGGAGCAGTGTGCT 2638

Db 2581 ACAAGCTCTCTCTGACACCCCTACCTTCTAATCAGGAGCAGTGAGCAGTGCTGCT 2640  
Qy 2639 CCTCTCTAGCACCTGCTTGTACCCAGCAAGGATAGGTTGGAAACCTATCATTTCC 2698  
Db 2641 CCTCTCTAGCACCCCTGCTTGTACCCAGCAAGGATAGGTTGGAAACCTATCATTTCC 2700  
Qy 2699 TGTCTCTCTCTTAAGAGGAAGACGCTCTCTGTAGAAACCAAGACTTCTTTTCTG 2758  
Db 2701 TGTCTCTCTCTTAAGAGGAAGACGCTCTCTGTAGAAACCAAGACTTCTTTTCTG 2760  
Qy 2759 GCTCTTTTCTTTACACCTTAAAGCCAGAAACCTTCAGAGTATTGTGTGGGTGTGTT 2818  
Db 2761 GCTCTTTTCTTTACACCTTAAAGCCAGAAACCTTCAGAGTATTGTGTGGGTGTGTT 2820  
Qy 2819 TTATATTAGCATTTGGGGATGGGGTGGGAGGGGTTATAGTTCATCAGGGTTTCTTAA 2878  
Db 2821 TTATATTAGCATTTGGGGATGGGGTGGGAGGGGTTATAGTTCATCAGGGTTTCTTAA 2880  
Qy 2879 GAAATTGCTAACAAAGCACTTTTGACAAATGCTATCCAGCAGGAAAGAAAGGATATA 2938  
Db 2881 GAAATTGCTAACAAAGCACTTTTGACAAATGCTATCCAGCAGGAAAGAAAGGATATA 2940  
Qy 2939 TAACTGTTTTAAACCTTTCTTGGGGAATCCAAATTATAGTTCCTTTGATTTAAACAA 2998  
Db 2941 TAACTGTTTTAAACCTTTCTTGGGGAATCCAAATTATAGTTCCTTTGATTTAAACAA 3000  
Qy 2999 GAACAGCAAGGTTGTCGACGGGTAGGATGTCTTAAAGATTTGTCCTTGAATAAT 3058  
Db 3001 GAACAGCAAGGTTGTCGACGGGTAGGATGTCTTAAAGATTTGTCCTTGAATAAT 3060  
Qy 3059 ATGCTTCCTGATCAAGAGTACGATGTTGTCGCAACAGGAGCAAACTTCTTTTAAAT 3118  
Db 3061 ATGCTTCCTGATCAAGAGTACGATGTTGTCGCAACAGGAGCAAACTTCTTTTAAAT 3120  
Qy 3119 TCCTTCCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATTACTACAAATCAGACAT 3178  
Db 3121 TCCTTCCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATTACTACAAATCAGACAT 3180  
Qy 3179 GGCAAAACAAATAGCTGTTGCTTCCATAAACAAAGTGCATTTTAAAGTGTCTGCT 3238  
Db 3181 GGCAAAACAAATAGCTGTTGCTTCCATAAACAAAGTGCATTTTAAAGTGTCTGCT 3240  
Qy 3239 TACTAAGTCTGTTTATTAATCTCTCTTTATCTATATGGAATAAAGAGGAGTCA 3298  
Db 3241 TACTAAGTCTGTTTATTAATCTCTCTTTATCTATATGGAATAAAGAGGAGTCA 3300  
Qy 3299 TGTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTCGATC 3358  
Db 3301 TGTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACTTCCCTGTCGATC 3360  
Qy 3359 CCTTCTGAGGTATGSCCATCAAGACTTTTAGGCCATTCTTGTATGGAACAGATCCCTG 3418  
Db 3361 CCTTCTGAGGTATGSCCATCAAGACTTTTAGGCCATTCTTGTATGGAACAGATCCCTG 3420  
Qy 3419 CCCTGACTGTCAGCTATCTGAAAGTGGATCAGATTAATACTGGATTAATGTAATCTG 3478  
Db 3421 CCCTGACTGTCAGCTATCTGAAAGTGGATCAGATTAATACTGGATTAATGTAATCTG 3480  
Qy 3479 TTTTGGTTGTTCTTATCAACCCACACAGAGTTCCTTAACTGCTTCAGTTATAGTAAC 3538  
Db 3481 TTTTGGTTGTTCTTATCAACCCACACAGAGTTCCTTAACTGCTTCAGTTATAGTAAC 3540  
Qy 3539 TGACTGTTATATTCAATCAGAAGCGCCATAAGTTCAGTTGAGTATTGATCCCTAGATAAG 3598  
Db 3541 TGACTGTTATATTCAATCAGAAGCGCCATAAGTTCAGTTGAGTATTGATCCCTAGATAAG 3600  
Qy 3599 AACATGCAAAATCAGCAGGAATCTGTATACAGGGTAAGCACCAGGGAACAATAAGATTTT 3658  
Db 3601 AACATGCAAAATCAGCAGGAATCTGTATACAGGGTAAGCACCAGGGAACAATAAGATTTT 3660  
Qy 3659 TATAGATATAATTTAATTTTGTAA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3715

Db 3661 TATAGATATAATTTAATTTTGGTAAATGGGTAAAGGAGACCAATTTTGGAGAGCAAGCA 3720  
Qy 3716 AA---TCTTTTAAAAATAGTATGAATGTGAATACTAGAAAAAGATTTAAAAAATAGTAT 3772  
Db 3721 AATCTCTCTTTTAAAAAATAGTATGAATGTGAATACTAGAAAAAGATTTAAGAAATAGTAT 3780  
Qy 3773 GAGTGTGAGTACTAGGAGGAT 3794  
Db 3781 GAGTGTGAGTACTAGGAGGAT 3802

RESULT 3  
US-09-023-655-928  
; Sequence 928, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 928:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2714 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g1311504  
US-09-023-655-928

Query Match 70.9%; Score 2689.6; DB 4; Length 2714;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 132 TCACCTCCGACACAGACACAGGCGGCACACAGGCTCCGACACACACTTCGCTCTCCC 191  
Db 19 TCACAGAAACACACAGACAGAGCGGCACACAGGCTCCGACACACACTTCGCTCTCCC 78  
Qy 192 GCGCGCTCACACCCCTCTTGCCCTTGAGCCCTTGCCGTTGACGCGCGCGCGAGCTGGA 251  
Db 79 GCGCGCTCACACCCCTCTTGCCCTTGAGCCCTTGCCGTTGACGCGCGCGCGAGCTGGA 138  
Qy 252 CGCCCTTCCCGGGCTCACTTTTGCAACGCTGACGCTGCCGCGAGTGGAGGTGGA 311



Db ||||| 139 CGCCCTCCCGGCTCACTTTGCAACGCTGACGGTGCAGTGGCGGTGGGA 198  
Qy ||||| 312 ACAGGGGGATCCTCCCGCTGGTCAAGCCCAAGCCAGGAGCCCGGGACCTCTC 371  
Db ||||| 199 ACAGGGGGATCCTCCCGCTGGTCAAGCCCAAGCCAGGAGCCCGGGACCTCTC 258  
Qy ||||| 372 GGCTGTCTCTCCCATGAGTCGGGATCGCAGCATCCCGCCACAGCCGCTCACCGCCTCCG 431  
Db ||||| 259 GGCTGTCTCTCCCATGAGTCGGGATCGCAGCATCCCGCCACAGCCGCTCACCGCCTCCG 318  
Qy ||||| 432 GGAGCGCTGGCTTGTACCCGAGAGCCCTTCCGGGACAGAGCTGTGACTTCCCGCCAG 491  
Db ||||| 319 GGAGCGCTGGCTTGTACCCGAGAGCCCTTCCGGGACAGAGCTGTGACTTCCCGCCAG 378  
Qy ||||| 492 TGCAGATTTCCGGACAGCTCTAGAACTCGCTCTAAAGACGAAACCGCCACAGCACTC 551  
Db ||||| 379 TGCAGATTTCCGGACAGCTCTAGAACTCGCTCTAAAGACGAAACCGCCACAGCACTC 438  
Qy ||||| 552 AAAGCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCCCTGAGCTTGGACCTTAGC 611  
Db ||||| 439 AAAGCCACTCGGAAGAGGCGAGCCCGGCAAGCCCGGCCCTGAGCTTGGACCTTAGC 498  
Qy ||||| 612 GGTGCGGGAGCACTGCGCGGCTTGCCTTCGCGGAGAGCTTCCCTTCACTCTC 671  
Db ||||| 499 GGTGCGGGAGCACTGCGCGGCTTGCCTTCGCGGAGAGCTTCCCTTCACTCTC 558  
Qy ||||| 672 AGCTTCGCTGGAGAGCCCCAGCCACCATTCAGCGCGAAGATACCTCCAGATAT 731  
Db ||||| 559 AGCTTCGCTGGAGAGCCCCAGCCACCATTCAGCGCGAAGATACCTCCAGATAT 618  
Qy ||||| 732 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTTCCAGTTATGCGGCGAGAC 791  
Db ||||| 619 GCCCTGGCTCAAGCCCAATATAGCCCTTCCCTCCAGTTTCCAGTTATGCGGCGAGAC 678  
Qy ||||| 792 ATACAGCTCGGAATACACCAAGAGATCATGAAACCCGCACTACACCAAGCTGACATGGA 851  
Db ||||| 679 ATACAGCTCGGAATACACCAAGAGATCATGAAACCCGCACTACACCAAGCTGACATGGA 738  
Qy ||||| 852 CTTGGGAGCACTGAGATCAGGCTACAGCCACACAGCTCCCTGCGCCAGCATCAGTACTT 911  
Db ||||| 739 CTTGGGAGCACTGAGATCAGGCTACAGCCACACAGCTCCCTGCGCCAGCATCAGTACTT 798  
Qy ||||| 912 CGTGAGGGCTACTCGAGCAACTAGCACTCAAGCTTTCCTGCTGTACCAAAATGCGAGG 971  
Db ||||| 799 CGTGAGGGCTACTCGAGCAACTAGCACTCAAGCTTTCCTGCTGTACCAAAATGCGAGG 858  
Qy ||||| 972 GCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCAACCATACCAACCA 1031  
Db ||||| 859 GCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCAACCATACCAACCA 918  
Qy ||||| 1032 CCACCAACCAACCATACAGCAGCAGCATCAGCAGCATCCATTCCTCCAGGCTC 1091  
Db ||||| 919 CCACCAACCAACCATACAGCAGCAGCATCAGCAGCATCCATTCCTCCAGGCTC 978  
Qy ||||| 1092 CAGCCCGAGGAGGAGTGTCTGCCAGCACCTTCAAGTGTACTTCAAGCAGTCCCGCCAGTCC 1151  
Db ||||| 979 CAGCCCGAGGAGGAGTGTCTGCCAGCACCTTCAAGTGTACTTCAAGCAGTCCCGCCAGTCC 1038  
Qy ||||| 1152 CACCCCAACAGCGCGCTTCCCGCCAGCAGCGGGCGGTATGGAGAGGAGCACTGCC 1211  
Db ||||| 1039 CACCCCAACAGCGCGCTTCCCGCCAGCAGCGGGCGGTATGGAGAGGAGCACTGCC 1098  
Qy ||||| 1212 CTGCGCGCCGCTGATCGACCCGCGCGCTGCTGGACCCCGCGGATGAAGGCGGTCC 1271  
Db ||||| 1099 CTGCGCGCCGCTGATCGACCCGCGCGCTGCTGGACCCCGCGGATGAAGGCGGTCC 1158  
Qy ||||| 1272 CAGGTGCGCGCGGCTTCCCGCTTCTTCACTTCAAGCCCTCGCGCGGCTATCCCG 1331  
Db ||||| 1159 CAGGTGCGCGCGGCTTCCCGCTTCTTCACTTCAAGCCCTCGCGCGGCTATCCCG 1218  
Qy ||||| 1332 CGCGCCAGCGCGCGGCGGCGACCATCTGGCTACGACCGGAGCGCGCTGCGGCT 1391

Db ||||| 1219 CGCGCCAGCCCGCGCGGCGCCACCACTTCGGCTACGACCCGAGCGCGCTGCGCGCT 1278  
Qy ||||| 1392 CAGCTTCGCGCTGGAGAGCCGAGCCGCGCGGAGCCAGCCAGCGCGCGCTTGAAGACCA 1451  
Db ||||| 1279 CAGCTTCGCGCTGGAGAGCCGAGCCGCGCGGAGCCAGCCAGCGCGCTTGAAGACCA 1338  
Qy ||||| 1452 CCCTGAGGGCTCGCTGCTGGCAAGAGGGCGCGCGCTTCCCGCTTCCCGCTTCCCGCT 1511  
Db ||||| 1339 CCCTGAGGGCTCGCTGCTGGCAAGAGGGCGCGCGCTTCCCGCTTCCCGCTTCCCGCT 1398  
Qy ||||| 1512 CAGCCCTCCCTTACCGCTTCAGCCCTGCTGGGAGAGTCCAGCCCTGCTGCGCGCC 1571  
Db ||||| 1399 CAGCCCTCCCTTACCGCTTCAGCCCTGCTGGGAGAGTCCAGCCCTGCTGCGCGCC 1458  
Qy ||||| 1572 CAGCAGAGCTCGCTGCTGGCGAGGGCACTGTGCGCTGTGGGGACAAACGCGCTG 1631  
Db ||||| 1459 CAGCAGAGCTCGCTGCTGGCGAGGGCACTGTGCGCTGTGGGGACAAACGCGCTG 1518  
Qy ||||| 1632 CAGCAGCTACGGCGTGGAAACCTGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1691  
Db ||||| 1519 CAGCAGCTACGGCGTGGAAACCTGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGCA 1578  
Qy ||||| 1692 GAAAGATGCAAAATATGTTTGCCTGGCAATATAAACTGCCAGTACAGAGACGCTG 1751  
Db ||||| 1579 GAAAGATGCAAAATATGTTTGCCTGGCAATATAAACTGCCAGTACAGAGACGCTG 1638  
Qy ||||| 1752 AAACCGATGTCACTGTCTGATTTTCAAGAGTGTCTCAGTGTGGAAATGGTAAAGAGT 1811  
Db ||||| 1639 AAACCGATGTCACTGTCTGATTTTCAAGAGTGTCTCAGTGTGGAAATGGTAAAGAGT 1698  
Qy ||||| 1812 TGTCCGTACAGATGTCTGAAAGGGAGGAGTGTCTGCTTCCAAACCAAGAGGCC 1871  
Db ||||| 1699 TGTCCGTACAGATGTCTGAAAGGGAGGAGTGTCTGCTTCCAAACCAAGAGGCC 1758  
Qy ||||| 1872 ATTACACAGGAACCTTCCAGCCCTTCCACCTTCTCCTCCATCTGCATGATGAATGC 1931  
Db ||||| 1759 ATTACACAGGAACCTTCCAGCCCTTCCACCTTCTCCTCCATCTGCATGATGAATGC 1818  
Qy ||||| 1932 CTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATATCTGCC 1991  
Db ||||| 1819 CTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTTCCAGATATCTGCC 1878  
Qy ||||| 1992 CACTGACAGGCTGTCTGAGGCAAGATGTGAGCATGTGCAACAATTTACACCTCT 2051  
Db ||||| 1879 CACTGACAGGCTGTCTGAGGCAAGATGTGAGCATGTGCAACAATTTACACCTCT 1938  
Qy ||||| 2052 GACAGCTTCCATTTGATGATCTCAGAGCTGCGGAGAAAGATTCGGGATTTACTGATCT 2111  
Db ||||| 1939 GACAGCTTCCATTTGATGATCTCAGAGCTGCGGAGAAAGATTCGGGATTTACTGATCT 1998  
Qy ||||| 2112 CCCAAAGAGATCAGACATTTACTTTTGAATCAGCCCTTTTGGAGCTGTTGCTCAG 2171  
Db ||||| 1999 CCCAAAGAGATCAGACATTTACTTTTGAATCAGCCCTTTTGGAGCTGTTGCTCAG 2058  
Qy ||||| 2172 ACTTTCATCAGGTCAAACTGCTGAAATAGTTTGTGTTCTGCAATGAGACTGCTCT 2231  
Db ||||| 2059 ACTTTCATCAGGTCAAACTGCTGAAATAGTTTGTGTTCTGCAATGAGACTGCTCT 2118  
Qy ||||| 2232 GCATCGACTTTCAGTGCCTTCTGATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2291  
Db ||||| 2119 GCATCGACTTTCAGTGCCTTCTGATTTGGGAGTGGCTCGACTCTATTAAGACTTTTC 2178  
Qy ||||| 2292 CTTAAATTTGAGAGCTGAACCTTTGATATCCAGCCCTTAGCCTGCTGAGCACTGAG 2351  
Db ||||| 2179 CTTAAATTTGAGAGCTGAACCTTTGATATCCAGCCCTTAGCCTGCTGAGCACTGAG 2238  
Qy ||||| 2352 CATGATCAGAAAGACATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAA 2411  
Db ||||| 2239 CATGATCAGAAAGACATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAA 2298  
Qy ||||| 2412 GATCAGAGCAGTTTAAAGAACCAAGAGTGAAGGAGCAGGCTCTGAGGCCACCGAGTC 2471  
Db ||||| 2299 GATCAGAGCAGTTTAAAGAACCAAGAGTGAAGGAGCAGGCTCTGAGGCCACCGAGTC 2358





CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/023.655  
 FILING DATE: HEREWITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1400:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3427 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9415822  
 US-09-023-655-1400

Query Match 13.1%; Score 497; DB 4; Length 3427;

Best Local Similarity 59.1%; Pred. No. 3.7e-105; Indels 144; Gaps 11;  
 Matches 1135; Conservative 0; Mismatches 640;

Qy	722	CTCCAGATATGCTTCCGATGCTTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATG	781
Db	310	CTGAAGCCATGCTTGTGTTTCAGGCGCAGTATGGTCTCTCGCTCAAGGAGCCAGCCCG	369
Qy	782	CGGCGCAGACATACAGCTC-----GGATACACACAGGAGATCATGAACCCCG	829
Db	370	CTTCTCAGAGCTCAGTTTACCACCTCTTCGGGAGATATACAGTCCGATTTCTTAATCCAG	429
Qy	830	ACTACCAAGCTGACCATGACCTTGGCAGCAGTACAGCTACAGGCTACAGCCACGCT	889
Db	430	AGTTTGTCAAGTTTAGCATGACCTTCAACACACTGAATCACT-----GCCACCACTT	483
Qy	890	CCCTGCCAGCATCAGTACCTTCTGTGGAGGCTACTCGAGCAACTTACGAACTCAAGCCTT	949
Db	484	CTCTCCCGAGCTTCAGTACCTTTATGGACAACCTACAGCAGGCTACAGCTCAAGCCAC	543
Qy	950	CTTGCTGTACAAATGACGGCCCTTTGATCAAGTGGAGAGGGCGGGCGCCAGCT	1009
Db	544	CTTGCTGTACAAATGCCCCCTGTCC-----GGACAGCAGTCTCTCCA	585
Qy	1010	ACCATCACCATCAC	1069
Db	586	TTAAGGTAGAGACATTGATGACAACTACAGCAACACAGCCACCTGCCCCCAGT	645
Qy	1070	AGCCATCAATTTCTCCAGCCTCCAGCCCGGAGGAGGAGTGTGTCACAGCACCTCCATGT	1129
Db	646	-----CTGAGGAGATGATCGGCACCTCCGGTCTGGTTT	678
Qy	1130	ACTTCAAGCAGTCCCGCCAGTCCAGCCCGCCAGCCCGCCCTTCCCCCGGAGGGGG	1189
Db	679	ACTAAGCCCTCTCTCGCCCGGAGCCACACACACACACACACACACACACACACACAC	738
Qy	1190	CGTTATGGAGCAGCAGTCCCTCGGCGCCGCTGATCGACCCCGCGCTGCTGG	1249

Db	739	CCATGTGGAGCA-----CCCGGATCTCTCCACNACTTCCACAGCAACTACGTGGCCA	792
Qy	1250	ACCCGCGATGAAGCGGTCCCCACGTTGGCGCGCGGTTCCTGCTTTCACATTCA	1309
Db	793	CTAGCGCATGATCGAGCAGAGGAAACGCGAGTCTCCCGCTCTTCCCTCTTCTCTTTA	852
Qy	1310	AGCCTCGCGCGCATCCCCCGCGCCAGCCCGCGCGCGCGCCGACACACCTCGGTAG	1369
Db	853	AGCAATGCCCCCTGGCACCCCCGGTGTCTAG-----TT	885
Qy	1370	ACCCGAGCGCGCTGCGCGCTCAGCCTGCGCTGGGAGCGCGACCGCCGCGGCGAGCC	1429
Db	886	GCCAGATGCGTTTCGACGGGCCCCGACGTCCCATGAAACCCGAGCGCCGCGGAGCC	945
Qy	1430	AGCGCGCGCGCTTTGAGAGCCACCGGTACGGGTGCGCTGGCCAAAGAGGCGGCGCCG	1489
Db	946	ACCACGTGGTGACGGGCGAGACCTTCGCTGTGCCAACCCCATTCGCAAGCCCGCTCCA	1005
Qy	1490	TGGCTTCCCGCTCTCGGCTCAGCCCTCCCTACCGCTCCAGCTCCAGCTGCGGCGAGA	1549
Db	1006	TGGCTTCCCGGCTTCAGATCGGCCACGCTCTCAGCTGCTCGACACGAGG-----	1059
Qy	1550	GTCCAGCCTGCGCTGCGCGCCAGCAGGAGCTCGTCTGTGGGAGGCGCAGCTGTGCGG	1609
Db	1060	-----TGCCCTCACCGCGCTCGCGGGCTCCCCCTCCAAAGAGGGGTGTGCGGTG	1110
Qy	1610	TGTGCGGAGCAACGCGCTTGCAGCAGCTACGGGTGCGAACTTCGAGGGCTCAAGG	1669
Db	1111	TGTGTGGGAGCAACGCGCTTGCACACTACGCGCTGCGCACTGTGAGGGCTGCAAG	1170
Qy	1670	GCTTTTTCAGAGAACAGTGCAGAAATGCAAAATATGTTTGCCTGCGCAATAAACT	1729
Db	1171	GCTTCTTTAAGCGCACAGTGCAGAAATGCAAAATAGTGTGTGTAGCAATAAAACT	1230
Qy	1730	GCCAGTAGACAGAGAGCTGCAAAACCGATGTAGTACTGTCTGATTTTCAGAAAGTCTCA	1789
Db	1231	GCCAGTAGACAGAGAGCTGCGCGGATCGCTGTAGTACTGTCTGATTTTCAGAAAGTCTCA	1290
Qy	1790	GTGTGGAATGGTAAAAAGTGTTCGTACAGATAGTCTGAAAGGGAGGAGAGTCTGTC	1849
Db	1291	CTGTGGATGCTCAAGAGAGTGTTCGACAGACAGTTTAAAGGCGCGGAGAGGTCTT	1350
Qy	1850	TGCTTCCAAACCAAGAGCCCATTAACAGGAACTTCTCAGCCCTCTCCACCTCTC	1909
Db	1351	TGCTTCCAAACCAAGAGCCCATTAACAGGAACTTCTCAGCCCTCTCCACCTCTC	1395
Qy	1910	CTCCAACTCGCATGATGAATGCCCTTGTCCGAGCTTTAAACAGACTCAACACC-----CA	1963
Db	1396	CCCGGTGAGTCTGATCAGTCCCTCTGTCAGGGCCCATGTGCTGACTCCAAACCGGCTATGA	1455
Qy	1964	GAGATCTTGATTTCCAGATACGTGTC---CACTGACAGGCTGCTCGAGGACAGATG	2020
Db	1456	CCAGCTGGACTATTCAGGTTCCAGGCGAACCTGACTATCAAAATGATGGAGATGACA	1515
Qy	2021	CTGAGCATGTGCACAAATTTCAACCTCTCGACAGCTCCATTTGATGTATCCAGAACT	2080
Db	1516	CCCAGCATATCCAGCAATTTCTATGATCTCTGAGTGGCTCCATGGAGATCATCCGGGCT	1575
Qy	2081	GGGAGAGAAAGATTCCGGGATTTTACTGATCTCCCAAGAAAGATCAGACATTTACTTTTG	2140
Db	1576	GGGAGAGAGATCTCTGGCTTCGAGACCTGCCCCNAAGCCGACCAAGACCTGCTTTTG	1635
Qy	2141	AATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTTCCATCAGGTCAAACACTGCTGAG	2200
Db	1636	AATCAGCCTTTCTAGAACTGTTGTCCTTCGATAGCATACAGGTCCAAACCCAGTGGAG	1695
Qy	2201	ATAGTTTGTGTTCTGCAATGGAATGCTCTGCTGATCGACTTTCAGTGCCTGCTGATTTG	2260
Db	1696	GTAAACTCATCTTTTGAATGGGTGGTCTTTCGACAGGTTGCAATGCTTCTGCTGCTTG	1755
Qy	2261	GGGAGTGTGCTGCTATTAAGAACTTTTCTTAAATTTTTCAGAGGCTGTAACCTTGATA	2320
Db	1756	GGGATGGATGATTCATTTGTTGATTTCTCTCCACTTGCAGATATGAACATCGACA	1815

Db	586	TTAAGGTAGAGACATT	CAGATGCAACAACTAC	CAGCAACAGCAGCACTGCCGCCCCAGT	645
Qy	1070	AGCCATCCATTCT	CCAGCCTCCAGCCCGGAGGAC	GAGGTGTGTCGCCAGCACTCCATGT	1129
Db	646	-----	-----	CTGAGGAGATGATCGGCACCTCCGGGTGGTTT	678
Qy	1130	ACTTCAAGCAGTCC	CCACCGTCCACCCCA	CCACCGCGGCTTCCCGCGAGCGGGG	1189
Db	679	ACTACAAGCCCT	CCTCGCCCCGAGCC	CACACCCCGGGTTCAGGTGCAACAGCC	738
Qy	1190	CGTTATGGACAGG	CAGCTGCCTCGGCGCCGGTGC	ATCGCAACCGCGCCCGCTGCTGG	1249
Db	739	CCATGTGGGACGA	-----	CCCGGATCTCTCCACAACTTCCACGAACTACGTGGCCA	792
Qy	1250	ACCCGCGGATGA	AGCGGTCCCAACGGTGGCGCGCGCGCTTCCGCTCTTCCACTTCA	1309	
Db	793	CTACGCACATGAT	TCGAGCAGAGGAAAA	CCCGAGTCTCCCGCTCTCCCTCTCTCTCTTTA	852
Qy	1310	AGCCCTCGCGCG	CATCCCCCGCGCCAGCCGCGCGCGCCACCACTCGGCTACG	1369	
Db	853	AGCAATGCGCC	CTGGCACCCCGGTCTAG	-----	TT 885
Qy	1370	ACCCGACGGCG	CTGCGCGCTCAGCTTCGCTGGAGCCGAGCGCGCGGGGAGCC	1429	
Db	886	GCCAGATGCGT	TCGAGCGGCCCTGCA	CGTCCCATGAACCGGAGCCCGCGCGAGCC	945
Qy	1430	AGGCGCGCGCT	TGAGGCCACCGTAGCGGTGCGCTGGCGAAGAGGGCGGCCCGC	1489	
Db	946	ACCACTGTGTGA	CGGGCAGACCTTCGTGTGCCAAACCCCATTCGCAACCGCGCTCCA	1005	
Qy	1490	TGGCTTCCCGCT	CTCGGCTCACGCGCTTCCCTACCGCTGCAGCTGTGGGCGAGA	1549	
Db	1006	TGGCTTCCCGG	CTCGAGTCGGCACGCGCTCTCAGCTGTCTGCAACGCAAG	-----	1059
Qy	1550	GTCCAGCTTGC	GTCCGCCAGCAGAGCTCGTCTGTGGCGAGGCGACGTGCGG	1609	
Db	1060	-----	TGCCCCCTACCCCGCTCGCGGGCTCCCCCTCAACAGAGGGCTGTGGCTG	1110	
Qy	1610	TGTGCGGGGCA	CACGCGCGCTGCAGCACTACGCGCTGCGAACCTCGAGGGGTGCAAGG	1669	
Db	1111	TGTGTGGGGA	CACGCGCTTCCACACTACGCGGTGCGCACTGTGAGGGGTGCAAG	1170	
Qy	1670	GCTTTTCAAGAG	AACAGTGCAGAAATGCAAAATATGTTGCTGGCAAAATAAAACT	1729	
Db	1171	GCTTCTTAAAG	CGCACAGTGCAGAAAAATGCAAAATACGTGTGTTAGCAAAATAAAACT	1230	
Qy	1730	GCCAGTAGACA	AGAGACGTGAAACCGATGTAGTACTTCGATTTTCAGAAAGTGTCTCA	1789	
Db	1231	GCCAGTGGACA	CGCTGCGCGAATCGCTGTCAGTACTCCGATTTTCAGAAAGTGCCTGG	1290	
Qy	1790	GTGTTGGAAT	GGTAAAAAGTGTTCGTTACAGATAGTCTGAAAGGAGGAGGTGCTC	1849	
Db	1291	CTGTTGGGAT	GGTCAAAAGATGTTTCGCACAGCTTTAAAGGCGGAGAGGTGCTT	1350	
Qy	1850	TGCTTCCAAAC	CAAGGCCATTACAAAGGAACCTTCTCAGCCCTCTCCACCTTCTC	1909	
Db	1351	TGCCCCGAA	ACCAGAGCCCA	-----	CAGGAGCCCTCTCCCCCTTCG 1395
Qy	1910	CTCCAAATCG	CATGAATGCCCTGTTCGAGCTTTAAACAGCTCAACACC	-----	CA 1963
Db	1396	CCCCGGT	CGAGTCTGATCAGTGCCTTCGTCAGGGCCCATGTGCACTCAACCGCGCTATGA	1455	
Qy	1964	GAGATCTTGAT	TATTCAGATACTGTCC	---	CACGTACAGGGTGTGCGAGGCACAGATG 2020
Db	1456	CCAGCTCGCAT	TATTCAGGTTCCAGGGCAACCTTGACTATCAATGAGTGGAGATGACA	1515	
Qy	2021	CTGAGCATGT	GCACAAATCTCAACCTCTGACAGCCTCATATGATGTATTCAGAGCT	2080	
Db	1516	CCAGCATAT	CCAGCAATTCCTATGATCTCCTGACTGGCTCCATGAGATCATCCGGGCT	1575	
Qy	2081	GGGCGAAAGAT	TCCGGATTTACTGATCTCCCCAAAGAGATCAGACATTTACTTTATG	2144	
Db	1576	GGGCGAGAGAT	CTCCGTTCGAGACCTGCCCCAAGCGGACCAAGCACTGCTTTTGT	1633	

QY	2321	TCCAAGCCTTAGGCTGCTGTACAGCTAGCATGATCA	CAGAAAGACATGGGTTAAAG	2380
DB	1816	TTTTCTGCTTCTCTCTGATTTGCTGCTGCTATGGT	CACAGAGAGACACGGGCTCAAGG	1875
QY	2381	AACCAAGAGAGTCGAAGAGCTATGCAACAGATCA	CAGCAGTTTAAAAGACCACCAGA	2440
DB	1876	AAACCAAGAGAGTGAAGAAGCTGAAAACAGATGT	AAATTGTCTCAAGACCAACGTGA	1935
QY	2441	GTAAGGACAGGCTCTGGAGCCCAACC-----	GAGTCAAGGTCCTGGGTCCCTCGG	2491
DB	1936	CTTTTCAACAATGGGGGTGGAACCGCCCAATTA	TTTGTCCAAACTGTTGGGGAAGTCC	1995
QY	2492	TAGAACTGAGGAAGATCTGCACCTCTGGGCTCC	AGCGCATCTTCTACTCTGAAGCTGGAAG	2551
DB	1996	CAGAACTTCGTACCTTTGGCACAGGGGCTACAG	CGCAATTTTCTACTCTGAATTTGGAAG	2055
QY	2552	ACTTGGTGTCTCCACCTTCCATCATTTGACAA	AGCTCTTCTCTGGACACCTTACTTTCTAA	2610
DB	2056	ACTTGGTGCCACCGCCAGCAATAATTGACAA	AACTTTTCTCTGGACACTTTTACTTTCTAA	2114
RESULT 7				
US-09-949-016-828				
; Sequence 828, Application US/09949016				
; Patent No. 6812339				
; GENERAL INFORMATION:				
; APPLICANT: VENTER, J. Craig et al.				
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED				
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF				
; FILE REFERENCE: CL001307				
; CURRENT APPLICATION NUMBER: US/09/949,016				
; CURRENT FILING DATE: 2000-04-14				
; PRIOR APPLICATION NUMBER: 60/241,755				
; PRIOR FILING DATE: 2000-10-20				
; PRIOR APPLICATION NUMBER: 60/237,768				
; PRIOR FILING DATE: 2000-10-03				
; PRIOR APPLICATION NUMBER: 60/231,498				
; PRIOR FILING DATE: 2000-09-08				
; NUMBER OF SEQ ID NOS: 207012				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 828				
; LENGTH: 3427				
; TYPE: DNA				
; ORGANISM: Human				
US-09-949-016-828				
Query Match				
Best Local Similarity 13.1%; Score 497; DB 4; Length 3427;				
Matches 1135; Conservative 0; Mismatches 640; Indels 144; Gaps 11;				
QY	722	CTCCAGATATGCGCTCGTCCAAAGCCCAATATAGCC	CTTCCCTCCAGGTTCCAGTTATG	781
DB	310	CTGAAGCCATGCTTGTGTTCAGGGCGAGTATGG	GTCTCGCCTCAAGGAGCCAGCCCG	369
QY	782	CGGCGCAGACATACAGCTC-----	GGAATACACCAAGGAGATCATGA	829
DB	370	CTTCTCAGAGCTACAGTTACCTCTTCGGGAGAA	TACAGTCCGATTTCTTAACCTCAG	429
QY	830	ACTACACCAAGCTGACATGGACCTTTGGGAGCA	CTGAGATCAGGCTCAGCCACCACGT	889
DB	430	AGTTTGTCAAGTTTAGCATGGACCTCACAACA	CTGAAATCACT-----GCCACCACTT	483
QY	890	CCCTGCCGACATCAGTACCTTCTGTGGAGGCT	TACTCGAGCAACTACGAACTCAAGCCTT	949
DB	484	CTCTCCCGAGCTTCAGTACCTTTATGAGAACT	TACGACAGGCTACGAGTCAAGCCAC	543
QY	950	CTCTGCGTGTACCAAAATGCAGCGCCCTTGAT	CAAAATGAGGAGGGCGGGCCCGCT	1009
DB	544	CTTGTCTGTACCAAAATGCCCTCTGTC-----	GGACAGCAGTCTCTCCA	585
QY	1010	ACCATCACTATCACCACCAACCAACCAACCA	CCACCACTATCACCAGCAGCATCAGC	1069



Qy 1964 GAGATCTTGATTATTCAGATACCTGTC---CACTGACCAAGCTGCTGCGAGCACAGATG 2020  
Db 1457 CCAGCTGGACTATTCAGAGTTCAGGCGAACCTGACTATCAATGAGTGGAGATGACA 1516  
Qy 2021 CTGAGATGTGCAACAATTTACAACTCTCGAGAGCTCCATTTGATGATATCCAGAAGCT 2080  
Db 1517 CCCAGCATATCCAGCAATTTCTATGATCTCTGACTGCTGCTGCTGCTGCTGCTGCTG 1576  
Qy 2081 GGGCAGAAAAGATTCGGGATTTACTGATCTCCCAAGAAAGATCAGACATTTACTTATTG 2140  
Db 1577 GGGCAGAGAAGATCCCTGGCTTCGAGACCTGCCAAGCGCACAGACCTGCTTTTGG 1636  
Qy 2141 AATCAGCCTTTTGGAGCTGTTTGTCTCTCAGACTTTCCATCAGGTCAAACTGCTGAAG 2200  
Db 1637 AATCAGCTTTCTTAGAATCTTTGCTCTCTGATAGCATACAGGTCCAAACCCAGTGGAG 1696  
Qy 2201 ATAGTTGTGTTCTGCAATGAGCTTGTCTGATGATGATGATGATGATGATGATGATG 2260  
Db 1697 GTAAACTCATCTTTTGAATGGGTGCTTCTGACAGGTTGCAATGCTGCTGCTGCTG 1756  
Qy 2261 GGGAGTGTCTGATCTATTAAGACTTTTCTTAAATTTGACAGCTGTAACCTTGATA 2320  
Db 1757 GGGAAATGATGATTCATTTGTTGATTTCTCTCACTTGCAGAAATATGAACATCGACA 1816  
Qy 2321 TCCAAGCCTTAGCCTGTGCTGACACTGAGCATGATCAGAAAGACATGGGTTAAAG 2380  
Db 1817 TTTCTGCTTCTCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876  
Qy 2381 AACCAGAGAGTGGAGAGTATGCAACAGATCAGACAGTAAAGACACACAGA 2440  
Db 1877 AACCAGAGAGTGGAGAACTGCAAAACAAGATTTGTTCTCAAGACACACGTGA 1936  
Qy 2441 GTAAGGACAGCTCTGAGGCCACCC-----GAGTCCAAGTCTGCTGGTGGCTCTGG 2491  
Db 1937 CTTCACAATGGGGGTGTAACCGCCCAATTTGTTCCAACTGTTGGGAGTCTC 1996  
Qy 2492 TAGAATCAGAGAAATCTGCACTTGGCTTCCAGCGCATCTTCTACCTGAAGCTGGAAG 2551  
Db 1997 CAGAATCTGCTACCTTTGCAACAGGGGCTACAGCGCATTTCTACCTGAAATTTGGAAG 2056  
Qy 2552 ACTTGGTCTCCACCTTCCATCATTTGACAGCTCTTCTGAGACACCTACCTTTCTAA 2610  
Db 2057 ACTTGGTCCACCGCCAGCAATAATTGACAACTTTTCTCGACACTTTACCTTTCTAA 2115

## RESULT 9

; Sequence 1, Application US/09277078  
; Patent No. 6312949  
; GENERAL INFORMATION:  
; APPLICANT: Sakurada, Kazuhiro  
; APPLICANT: Palmer, Theo  
; APPLICANT: Gage, Fred H.  
; TITLE OF INVENTION: REGULATION OF TYROSINE HYDROXYLASE  
; FILE REFERENCE: 07251/031001  
; CURRENT APPLICATION NUMBER: US/09/277,078  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1797  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)....(900)  
US-09-277-078-1

Query Match 12.8%; Score 486.4; DB 3; Length 1797;  
Best Local Similarity 58.1%; Pred. No. 7,9e-103;  
Matches 1105; Conservative 0; Mismatches 671; Indels 126; Gaps 9;

Qy 730 ATCCCTCGCTCCAAAGCCCAATATAGCCCTTCCCTTCCAGTTTCCAGTTATGCGGCGCAG 789  
Db 1 ATGCCTTGTTTCCAGGCGAGTATGGTCTCTCGCTCAAGAGCCAGCCCGCTTCTCAG 60  
Qy 790 ACATACAGCT-----CGGAATACACACCGAGATCATGAACCCCGACTACACC 837  
Db 61 AGTACAGTTTACCACTCTTCGGGAGAAATACAGTCCGATTTCTTAACTCCAGAGTTTGTG 120  
Qy 838 AAGCTGACATGAGACCTTGGCAGCAGTACAGATCACGGCTACAGCCACACAGTCCCTGCC 897  
Db 121 AAGTTAGCATGAGACCTTCCACACACTGAAAT-----TACTGCCACCACTTCTCTCCC 174  
Qy 898 AGCATCAGTACCTTCTGAGGGCTTACTCGAGCAACTACGAACTCAAGACCTTCTCTGCTG 957  
Db 175 AGCTTCAGTACCTTTATGGACAACCTACAGCACAGGCTTACGACGTCAAGCCACCTTGTG 234  
Qy 958 TACCAATGACGGGCTTGTGATCAAGTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1017  
Db 235 TACCAATGACGGGCTTGTGATCAAGTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 276  
Qy 1018 CATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1077  
Db 277 GAAGACATTCAGATGACAACTTACAGCAACACAGCCACCTGCCCCCTCAGT----- 328  
Qy 1078 ATTCTCCAGCCTTCCAGCCCGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1137  
Db 329 -----CCGAGGAGATGATGCCACACAGCGGGTGGTTTACTACAAAG 369  
Qy 1138 CAGTCCCCACCTGTCACCCCGGCTTCCCGGCGGAGGGGGGGGGGGGGGGGGGGGGGG 1197  
Db 370 CCTTTCGGCCCGGACACCCCGGCTTCCAGGTGAGCATAGCCCGATGAGTGG 429  
Qy 1198 GAGGAGCACTGCTCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1257  
Db 430 GACGA-----TCGGGCTCTTCAAACTTCCACAGAACTACGTGGCACTACGCAT 483  
Qy 1258 ATGAAGCGGTCCGACAGGTGGCGGGCGGCTTCCCGCTCTTCACTTCAAGCCCTCG 1317  
Db 484 ATGATCAGAGAGGAGACACCTGCTCCCGCTTTCACCTCTTCTCTTAAAGCAGTGG 543  
Qy 1318 CCGCGCATCCCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1377  
Db 544 CCGCGGCACTCTCTGTGTAG-----CTGCCAGATG 576  
Qy 1378 GCGCTGCGGCTCAGCTTCCGCTGGAGCGCAGCGCGGCGGCGGCGGCGGCGGCGGCGG 1437  
Db 577 CGCTTTGACGGGCTCTGACGCTCCCATGAACCGGAGCCCGGCGGCGGCGGCGGCGGTA 636  
Qy 1438 GCGTTGAGAGCCACCGGTACGGGCTCCGCTGGCCAAAGAGGGCGGCGGCGGCGGCTTC 1497  
Db 637 CTGGATGGGAGACCTTTCGCGGTGCCAATCCCATTCGCAAGCGCGGATCCATGGGCTTC 696  
Qy 1498 CCGCTCTCGGCTCAGCCCTTCCCTTCCCGCTTCCAGCTTCTGCGGCGAGAGTCCGAGC 1557  
Db 697 CCGGCTTGCAGATCGGCGGCGGCTGCGAGTGTTCACACGCGAG-----742  
Qy 1558 CTGCGCTGCGCGCCAGCAGAGCTGCTGCTGGCGAGGGCGGCGGCGGCGGCGGCGG 1617  
Db 743 -TGCCCTCGCGCGGCTTCCCGGCTTCCCTTCCAAATGAGGGTCTGCTGCTGCTGCTG 801  
Qy 1618 GACAGCGCGCTTCCAGCACTAGCGGTGGAGCTGCGAGGCTGCGAGGCTGCGAGGCTTTTC 1677  
Db 802 GACAGCGCGCTTCCAGCACTAGCGGTGGAGCTTGGAGGCTGCGAGGCTGCGAGGCTTTTC 861  
Qy 1678 AAGAGAACTGCAAGAAATATGCTTCTGCTGCGCAATAAATAAATAAATAAATAAATAA 1737  
Db 862 AAGCGCAGGTGCAAGAAATATGCTTCTGCTGCGCAATAAATAAATAAATAAATAAATA 921  
Qy 1738 GACAGAGAGCTGCAAGCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1797  
Db 922 GATAAGCGCGCGGAAATCGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981





2457 GGAGCCACCGAGTCCAAAGTCTCGTGGTCCCTGGTAGAAGTGAAGAGTCTGCACCT 2516  
1774 AGCCAGTGCCTGTACGCTGTGTTGGGCAAACTGCCGAGCTGGGACCTGTGACCCA 1833  
2517 GGGCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAT 2576  
1834 GGGCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGTGTCTCCACCTTCCATCAT 1893  
2577 TGACAAGCTCTTCTGGACACCTTCTTAATCAGGAGCAGTGGAGCAGTGC 2633  
1894 TGACAAGATCTTCTGACACGCTGCTTCTGACCCCTGCTGGGAACAGTGTGC 1950

RESULT 11  
US-09-566-921-58  
; Sequence 58, Application US/09566921  
; Patent No. 6682888  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debra W.  
; APPLICANT: Edwards, Carla M.  
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE  
; FILE REFERENCE: PA-0024 US  
; CURRENT APPLICATION NUMBER: US/09/566,921  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PERL Program  
; SEQ ID NO 58  
; LENGTH: 2637  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6682888 127112.18  
US-09-566-921-58

Query Match 9.0%; Score 340.6; DB 4; Length 2637;  
Best Local Similarity 61.3%; Pred. No. 7.1e-69;  
Matches 663; Conservative 0; Mismatches 389; Indels 29; Gaps 6;

1566 GCGCCCGAGGAGCTGCTGCTGCGGAGGCGACGTGCGGTGGGGGAGCAAGC 1625  
1030 GCGCCCGAGGCGGCGCCAGGTGGAAGTGAAGCCGCTGTGTGTGGGAGCAAGC 1089  
1626 CGCTCGCCAGCTACCGCTGCGAACCTGCGAGGCGTCAAGGGCTTTTCAAGAGAAC 1685  
1090 TTATGCCAGATATGCTGTCGCAATGTGAGGCTGCAAGGGCTTCTTCAAGCGCAC 1149  
1686 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1745  
1150 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1209  
1746 AGTGCAGAAAATGCAAAAATATGTTGCTGGCAATATAAATGCGCCAGTAGACAAGAG 1805  
1210 GCGCGAAGCGCTGCGCAATGTTGCTGCGCAATGTGAGGCTGCAAGGGCTTCTTCAAGCGCAC 1269  
1806 AGAAGTGTCTCGTACAGATAGTCTGAAGGGAGGAGAGTGTGCTGCTGCTTCCAAACCAA 1865  
1270 GGAAGTGTCTCGAAGACAGACCTGGAAGGGCGGCGGCGGCTACCTTCAAAACCCAA 1329  
1866 GAGCCCAATTAACAAGAACCTTCTGAGCCCTCTCCACCTTCTCCATCTGCATGAT 1925  
1330 GCAGCCCCCAGA-----TGCCTCCCTCCCAATCTCTCACTTCCCTGGTCCGT 1378  
1926 GAATGCCCTTCTCGAGCTTTAACAAGACTCAACACCCAGAGATCTTGATTTATCCAGATA 1985  
1379 GGAACACTTGACTCAGGGGCGGCGGCGGCGGCTGCAACTGCAACTCTCAAGTTCAGGAGC 1438  
1986 CTGTCCCACTGACCCAGGCTGTGAGGAGCAGAGATGTGAGCATGTGCAAAATCTACAA 2045  
1439 TGGTCTGCCCACTTTTGGGAAGGA-----AGATGCTGGGATGTACAGAGTCTTACGA 1493  
2046 CCTCTGACAGCTCCATTTGATGTATCCAGAAGCTGGGCAAGAAATTCGGGATTTAC 2105

1494 CTTGCTCTCCGCTTCTCTGGAGGTCAATCCGCAAGTGGGCGGAGAGATCCCTTGGCTTGC 1553  
2106 TGAATCTCCCAAGAGATCAGACATTAATTAATCAATCAGCTTTTGGAGCTTGTGT 2165  
1554 TGAATCTCACCAGGTGAGCAGGACCTGTGTGAGTGGGCTTCTTGGAGCTTGTGT 1613  
2166 CTTCAAGCTTTC-CATCAGGTCAAAACATGCTGTAAGATAAAGTGTGTGTCTGCAATGAC 2224  
1614 CTTCCGCTGGGTACAAGGTCTAAGCCAGGCGAGGCAAGCTCATCTTCTGCTCAGGCC 1673  
2225 TTGCTCTGATC-GACTTCAGTGCCTTCTGAGATTTGGGAGTGGCTCAGCTCTATATAA 2283  
1674 TGGTGTACACCGGGCTGCAGTGTGCTTCCGGGACTTGAATTCACAGTATCCTG 1733  
2284 GACTTTTCTTAATTTTGCAGAGCTTGAACCT-TGATATCCAAGCTTACCTTGCCTGTC 2342  
1734 GCTTCTCAAGGTCCCTGCACAGCTTGTGTCGATGTCCCTGCTTGCCTTGCCTGCTCTC 1793  
2343 AGCACTGAGCATGATCAGAGAAAGACATGGGTAAAGAACCAAGAGAGTGAAGAGCT 2402  
1794 TGGCTTGTCTCATCACCAGCCGATGGCTGCAAGGAGCCGCGGGGTGGAGGCT 1853  
2403 ATGCAACAGATCACAAGCAGTGTAAAGACACCAAGAGTAAGGAGACAGCTCTGGAGCC 2462  
1854 GCAGAACCGCATCGCCAGCTGCTGAGGAGCAGCTGGGAGCTGTGGGCGGAGCC 1913  
2463 -----CACCGAGTCCAAGTCTCTGGTCCCTGTTAGAACCTGAGGAGAGTGTGCA 2512  
1914 AGCCAGCAGCTGCTGCTGCTGCTGTTGGGCAAACTGCGGAGCTGCGGAGCTGTGCA 1973  
2513 CCTGGGCTCCAGCGATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCA 2572  
1974 CCCAGGCTGCGAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCA 2033  
2573 TCATTGACAGATCTTCTGACACAGCTGCGCTTCTGACCCCTGCTGCGGAGAGTGTG 2632  
2034 TCATTGACAGATCTTCTGACACAGCTGCGCTTCTGACCCCTGCTGCGGAGAGTGTG 2093  
2633 C 2633  
2094 C 2094

RESULT 12  
US-09-949-016-35444  
; Sequence 35444, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35444  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-35444

Query Match 8.3%; Score 315.6; DB 4; Length 601;  
Best Local Similarity 98.8%; Pred. No. 2.2e-63;  
Matches 328; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3466 TTACATGTAACGTTTGGTTGTTCTATCAACCCACAGTTCCCTAACTTGCTT 3525  
Db 1 TTACATGTAACGTTTGGTTGTTCTATCAACCCACAGTTCCCTAACTTGCTT 60  
QY 3526 CAGTTATAGTAACGTTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGAGTATTG 3585  
Db 61 CAGTTATAGTAACGTTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGAGTATTG 120  
QY 3586 ATCCCTAGATAAGAACATGCAATCAGCAGGAACCTGGTCATACAGGGTAAGCACCAGGA 3645  
Db 121 ATCCCTAGATAAGAACATGCAATCAGCAGGAACCTGGTCATACAGGGTAAGCACCAGGA 180  
QY 3646 CAATAAGGATTTTATAGATATAATTTTAAATTTTGTATTGTTAAGGAGACAATTTGG 3705  
Db 181 CAATAAGGATTTTATAGATATAATTTTAAATTTTGTATTGTTAAGGAGACAATTTGG 240  
QY 3706 AGAGCAAGCAAA--TCTTTTAAATAATAGTATGAATGTGAATCTAGAAAAGATTTAA 3762  
Db 241 AGAGCAAGCAAAATCTCTTTTAAATAATAGTATGAATGTGAATCTAGAAAAGATTTAA 300  
QY 3763 AAAATAGTATGAGTGTAGTACTAGGAAGGAT 3794  
Db 301 AAAATAGTATGAGTGTAGTACTAGGAAGGAT 332

## RESULT 13

US-09-949-016-135237  
; Sequence 135237, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 135237  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-135237

Query Match 8.3%; Score 315.6; DB 4; Length 601;  
Best Local Similarity 98.8%; Pred. No. 2.2e-63;  
Matches 328; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 3466 TTACATGTAACGTTTGGTTGTTCTATCAACCCACAGTTCCCTAACTTGCTT 3525  
Db 1 TTACATGTAACGTTTGGTTGTTCTATCAACCCACAGTTCCCTAACTTGCTT 60  
QY 3526 CAGTTATAGTAACGTTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGAGTATTG 3585  
Db 61 CAGTTATAGTAACGTTGTTATATTCATTCAAGAGCGCCATAAGTCAGTTGAGTATTG 120  
QY 3586 ATCCCTAGATAAGAACATGCAATCAGCAGGAACCTGGTCATACAGGGTAAGCACCAGGA 3645  
Db 121 ATCCCTAGATAAGAACATGCAATCAGCAGGAACCTGGTCATACAGGGTAAGCACCAGGA 180  
QY 3646 CAATAAGGATTTTATAGATATAATTTTAAATTTTGTATTGTTAAGGAGACAATTTGG 3705  
Db 181 CAATAAGGATTTTATAGATATAATTTTAAATTTTGTATTGTTAAGGAGACAATTTGG 240  
QY 3706 AGAGCAAGCAAA--TCTTTTAAATAATAGTATGAATGTGAATCTAGAAAAGATTTAA 3762  
Db 241 AGAGCAAGCAAAATCTCTTTTAAATAATAGTATGAATGTGAATCTAGAAAAGATTTAA 300

QY 3763 AAAATAGTATGAGTGTAGTACTAGGAAGGAT 3794  
Db 301 AAAATAGTATGAGTGTAGTACTAGGAAGGAT 332

## RESULT 14

US-09-949-016-135236  
; Sequence 135236, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 135236  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-135236

Query Match 6.2%; Score 236; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 7e-45;  
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2360 CAGAAGACATGGGTTAAAGAACCAAGAGAGTCAAGAGCTATGCAACAAGATCACAA 2419  
Db 366 CAGAAGACATGGGTTAAAGAACCAAGAGAGTCAAGAGCTATGCAACAAGATCACAA 425  
QY 2420 GCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAAGGTC 2479  
Db 426 GCAGTTTAAAGACCAACAGAGTAAAGGACAGGCTCTGGAGCCCAAGGTC 485  
QY 2480 TGGTGCCCTGGTAGAATCTGAGGAATCTGCACCTTCCATCATTTGCAAGCTCTTCTACC 2539  
Db 486 TGGTGCCCTGGTAGAATCTGAGGAATCTGCACCTTCCATCATTTGCAAGCTCTTCTACC 545  
QY 2540 TGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCTGGAC 2595  
Db 546 TGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCTGGAC 601

## RESULT 15

US-09-949-016-35441  
; Sequence 35441, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35441  
; LENGTH: 601

Tue Mar 22 09:54:21 2005

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-35441

Query Match      4.6%; Score 174.6; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 1.3e-30;
Matches 174; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1809 AGTTGTCCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCTTCCAAACCAAAGAG 1868
Db      |||:|||||
QY 178  AGTTGTCCGTACAGATAGTCTGAAAGGAGGAGAGGTCGTCTGCTTCCAAACCAAAGAG 237
Db      |||:|||||
QY 1869 CCCATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCAAATCTGCATGATGAA 1928
Db      |||:|||||
QY 238  CCCATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCAAATCTGCATGATGAA 297
Db      |||:|||||
QY 1929 TGCCCTTGTCCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTATTCCAGA 1983
Db      |||:|||||
QY 298  TGCYCTTGTCCGAGCTTTTAAACAGACTCAACACCCAGAGATCTTGATTATTCCAGA 352
Db      |||:|||||
```

Search completed: March 21, 2005, 20:50:42  
Job time : 606 secs





```
QY 2221 GGACTGTGCTGCATCGACTTCAGTGCCTTCGTGGATTGGGGAGTGGCTCGACTCTATT 2280
Db 2221 GGACTGTGCTGCATCGACTTCAGTGCCTTCGTGGATTGGGGAGTGGCTCGACTCTATT 2280
QY 2281 AAGACTTTTCCCTTAAATTTGAGAGCCTGAACCTTTGATATCAAGCCCTTAGCCTGCCTG 2340
Db 2281 AAGACTTTTCCCTTAAATTTGAGAGCCTGAACCTTTGATATCAAGCCCTTAGCCTGCCTG 2340
QY 2341 TCAGCACTGAGCATGATCACAGAAAGACATGGGTTAAAAGAACCAAGAGAGTCAAGAG 2400
Db 2341 TCAGCACTGAGCATGATCACAGAAAGACATGGGTTAAAAGAACCAAGAGAGTCAAGAG 2400
QY 2401 CTATCAACAAGATCACAGAGCAGTTTAAAAGACCACAGAGTAAGGGACAGGCTCTGGAG 2460
Db 2401 CTATCAACAAGATCACAGAGCAGTTTAAAAGACCACAGAGTAAGGGACAGGCTCTGGAG 2460
QY 2461 CCCACCGAGTCCAAAGTCTCGGTGCCCTGTGGTAGAACTCTGCAACCCCTGGGC 2520
Db 2461 CCCACCGAGTCCAAAGTCTCGGTGCCCTGTGGTAGAACTCTGCAACCCCTGGGC 2520
QY 2521 CTCAGAGCGATCTTCTACCTGAAGCTGGAAGACTTTGGTGTCTCCACCTTCATCATTCAC 2580
Db 2521 CTCAGAGCGATCTTCTACCTGAAGCTGGAAGACTTTGGTGTCTCCACCTTCATCATTCAC 2580
QY 2581 AAGCTCTTCTCGACACCTACCTTTCTAATCAGGAGCAGTGGAGCAGTGAGCTGCCTCC 2640
Db 2581 AAGCTCTTCTCGACACCTACCTTTCTAATCAGGAGCAGTGGAGCAGTGAGCTGCCTCC 2640
QY 2641 TCTCTAGCACCTGCTGTCTACGACGAAAGGATAGTTTGGAAACCTATCATTTCTCTG 2700
Db 2641 TCTCTAGCACCTGCTGTCTACGACGAAAGGATAGTTTGGAAACCTATCATTTCTCTG 2700
QY 2701 TCTTCTCTTAAGAGAAAGCAGCTCTGTAGAAAGCAAGACTTTCTTTTCTTTCTGGC 2760
Db 2701 TCTTCTCTTAAGAGAAAGCAGCTCTGTAGAAAGCAAGACTTTCTTTTCTTTCTGGC 2760
QY 2761 TCTTTTCTCTTACAACTTAAAGCCAGAACTTGCAGAGTATTGTGTGGGGTGTGTTT 2820
Db 2761 TCTTTTCTCTTACAACTTAAAGCCAGAACTTGCAGAGTATTGTGTGGGGTGTGTTT 2820
QY 2821 ATATTAGGCATTGGGGATGGGGTGGAGGGGGTTATAGTTTCATGAGGGTTTCTAAGA 2880
Db 2821 ATATTAGGCATTGGGGATGGGGTGGAGGGGGTTATAGTTTCATGAGGGTTTCTAAGA 2880
QY 2881 AATTCCTAACAAAGCATTGTTGGACAATGCTATCCAGCAGAGAAAGAAAGATTAATA 2940
Db 2881 AATTCCTAACAAAGCATTGTTGGACAATGCTATCCAGCAGAGAAAGAAAGATTAATA 2940
QY 2941 ACTGTTTTAAACTCTTTCTGGGAATCCAATTATAGTTGCTTCTATTAAAGCAAGA 3000
Db 2941 ACTGTTTTAAACTCTTTCTGGGAATCCAATTATAGTTGCTTCTATTAAAGCAAGA 3000
QY 3001 ACAGCAAGGGTGTGTCGCGAGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAGATAT 3060
Db 3001 ACAGCAAGGGTGTGTCGCGAGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAGATAT 3060
QY 3061 GCTTCTGTATCAAGAGTACGTATGTGTGCAAAACAAGCAGAACTTCTCTTTTAATTTC 3120
Db 3061 GCTTCTGTATCAAGAGTACGTATGTGTGCAAAACAAGCAGAACTTCTCTTTTAATTTC 3120
QY 3121 CTTCTTCTTTTATTTTAACTGTGAAGATGAGGATTAACCTACAAATCAGACATGG 3180
Db 3121 CTTCTTCTTTTATTTTAACTGTGAAGATGAGGATTAACCTACAAATCAGACATGG 3180
QY 3181 CAAAAACAATATGGCTGTGTTGCTTCCATAAACAAGTGAATTTTTTAAAGTCTGCTTTA 3240
Db 3181 CAAAAACAATATGGCTGTGTTGCTTCCATAAACAAGTGAATTTTTTAAAGTCTGCTTTA 3240
QY 3241 CTAAGCTGTGTTTATAACTCTCTCTTTTATTTATATGGAATAAAGAGAGGAGCAGTCATG 3300
Db 3241 CTAAGCTGTGTTTATAACTCTCTCTTTTATTTCTATATGGAATAAAGAGAGGAGCAGTCATG 3300
QY 3301 TTAGCAAAATGACACGTTTAATATCCCTAGCAGAGGCTGTGTTTCACTTCCCTGCTCATCC 3360
```

```
Db 3301 TTAGCAAAATGACACGTTTAATATCCCTAGCAGAGGCTGTGTTTCACTTCCCTGCTCATCC 3360
QY 3361 TTTGAGGTATGGCCCATCCAGACTTTTAGGCCAATTTTATGATGAACCCAGATCCCTGCC 3420
Db 3361 TTTGAGGTATGGCCCATCCAGACTTTTAGGCCAATTTTATGATGAACCCAGATCCCTGCC 3420
QY 3421 CTGACTGTCAGCTATCTCTGAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGT 3480
Db 3421 CTGACTGTCAGCTATCTCTGAAGTGGATCAGATTATAAAGTGGATTACATGTAAGTGT 3480
QY 3481 TTGGTTGTGTTCTATCAACCCACAGAGTTCCCTTAAACTTTCAGTTATAGTAACTG 3540
Db 3481 TTGGTTGTGTTCTATCAACCCACAGAGTTCCCTTAAACTTTCAGTTATAGTAACTG 3540
QY 3541 ACTGTTATTTTCAATTCAGAAAGCGCCATTAAGTCAGTTGATTTTGCATCCCTAGATAAGAA 3600
Db 3541 ACTGTTATTTTCAATTCAGAAAGCGCCATTAAGTCAGTTGATTTTGCATCCCTAGATAAGAA 3600
QY 3601 CATGCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCAGGAGGAGCAATTAAGGATTTTA 3660
Db 3601 CATGCAAAATCAGCAGGAACTGGTCATACAGGGTAAGCAGGAGGAGCAATTAAGGATTTTA 3660
QY 3661 TAGATATATTTAATTTTGTGTTTATGAGGAGCAATTTTGGAGAGCAAGCAAACTCT 3720
Db 3661 TAGATATATTTAATTTTGTGTTTATGAGGAGCAATTTTGGAGAGCAAGCAAACTCT 3720
QY 3721 TTTTAAAAAATAGTATGTAATGTAATCTAGAAAAAGATTTAAAAAATAGTATGAGTGTGA 3780
Db 3721 TTTTAAAAAATAGTATGTAATGTAATCTAGAAAAAGATTTAAAAAATAGTATGAGTGTGA 3780
QY 3781 GTACTAGGAAGGAT 3794
Db 3781 GTACTAGGAAGGAT 3794
```

## RESULT 2

```
US-09-853-386-46
; Sequence 46, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Breenihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972U1
; CURRENT APPLICATION NUMBER: US/09/853,386
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 3802
; TYPE: DNA
; ORGANISM: HUMAN
US-09-853-386-46
```

```
Query Match 98.1%; Score 3722; DB 9; Length 3802;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;
```

```
QY 1 ATAATGACGTGCCGAGAGAGCGGAAACGCGAGCGGGAGAGCGGAGTCTCTGCTCT 60
Db 1 ATRAATGACGTGCCGAGAGAGCGGAAACGCGAGCGGGAGAGCGGAGTCTCTGCTCT 60
QY 61 CCGCGCCCCACCCCTCCAGTCTCTGCTCTCTCTCCCTCCCATACACAGAGCGGCTCA 120
Db 61 CCGCGCCCCACCCCTCCAGTCTCTGCTCTCTCTCCCTCCCATACACAGAGCGGCTCA 120
QY 121 CACCGCTCCCTCACTCCGACACACAGACACAGCGGCGACACAGGCTCCG--CACACAC 178
```



121	Db	 CACCGCTCCTCACTCGAAACACAGACACAAGCGCGACACAGGCTCCGACACACAC 180
179	Qy	 ACTTCGCTCTCCGCGGCTCACACCCCTCTTGCCCTTGAGCCCTTCGCGTGAGCGCG 238
181	Db	 ACTTCGCTCTCCGCGGCTCACACCCCTCTTGCCCTTGAGCCCTTCGCGTGAGCGCG 240
239	Qy	 CGCGCAGCTGGACGCGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGGCAGTGCG 298
241	Db	 CGCGCAGCTGGACGCGCCCTCCCGGGCTCACTTTGCAACGCTGACGCTGCGGCAGTGCG 300
299	Qy	 CGTGGAGGTGGAAACAGCGGGGATCTCTCCCTCTGCTGTCACAGCCCAAGCCAGGACGCC 358
301	Db	 CGTGGAGGTGGAAACAGCGGGGATCTCTCCCTCTGCTGTCACAGCCCAAGCCAGGACGCC 360
359	Qy	 CGCGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCCACACGCG 418
361	Db	 CGCGGAACCTCTCGGCTGTGCTCTCCCATGAGTCGGGATCGCAGCATCCCCACACGCG 420
419	Qy	 CTCACCGCTCCGGAGCGCTGGCTGTGTACACCGCAGCCCTTCGCGGACAGCAGCTGT 478
421	Db	 CTCACCGCTCCGGAGCGCTGGCTGTGTACACCGCAGCCCTTCGCGGACAGCAGCTGT 480
479	Qy	 GACTCCCCCGCAGTCAGATTTCCGGAGACGCTCTCTAGAACTCGCTCTAAAGACGAAC 538
481	Db	 GACTCCCCCGCAGTCAGATTTCCGGAGACGCTCTCTAGAACTCGCTCTAAAGACGAAC 540
539	Qy	 CGCCACAGCACTCAAAGCCCACTCGCGAAGAGGCGAGCCCGCAAGCCCGGCGCTTGAGC 598
541	Db	 CGCCACAGCACTCAAAGCCCACTCGCGAAGAGGCGAGCCCGCAAGCCCGGCGCTTGAGC 600
599	Qy	 CTGAGACCTTTAGCGTCCGGGACGACTCGCGCGCTTCGCTCGCGGACGCTCCGCTC 658
601	Db	 CTGAGACCTTTAGCGTCCGGGACGACTCGCGCGCTTCGCTCGCGGACGCTCCGCTC 660
659	Qy	 CTCCTCACACTCTCAGCCTTCGCTGGAGAGACCCCGACCCCACTTCAGCGCGCAAGAT 718
661	Db	 CTCCTCACACTCTCAGCCTTCGCTGGAGAGACCCCGACCCCACTTCAGCGCGCAAGAT 720
719	Qy	 ACCTTCAGATATCGCTCGGTCAAGCCCAATATAGCCCTTCGCTTCGAGGTTCCAGTT 778
721	Db	 ACCTTCAGATATCGCTCGGTCAAGCCCAATATAGCCCTTCGCTTCGAGGTTCCAGTT 780
779	Qy	 ATCGGGCGCAGCATACAGCTCGGAATACACACGAGATCATGAACCCCGACTTACACCA 838
781	Db	 ATCGGGCGCAGCATACAGCTCGGAATACACACGAGATCATGAACCCCGACTTACACCA 840
839	Qy	 AGCTGACATGGACCTTTGGGAGCACTAGATACGGCTACAGCCACACAGCTCCCTGCCCA 898
841	Db	 AGCTGACATGGACCTTTGGGAGCACTAGATACGGCTACAGCCACACAGCTCCCTGCCCA 900
899	Qy	 GCATCAGTACCTTCGTGGAGGGCTACTCGAGCAACTACGAACCTTCCTGCGGTGT 958
901	Db	 GCATCAGTACCTTCGTGGAGGGCTACTCGAGCAACTACGAACCTTCCTGCGGTGT 960
959	Qy	 ACCAAAATCAGCGGCCCTTGATCAAAAGTGGAGGAGGGCGCGCCAGGTACCATCACC 1018
961	Db	 ACCAAAATCAGCGGCCCTTGATCAAAAGTGGAGGAGGGCGCGCCAGGTACCATCACC 1020
1019	Qy	 ATCACCAACACACACACACACACACCATACACAGCAGCAGCATCAGCAGCATCCCA 1078
1021	Db	 ATCACCAACACACACACACACACCATACACAGCAGCAGCATCAGCAGCATCCCA 1080
1079	Qy	 TTCTCTCAGCTCCAGCCCGGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCGAAGC 1138
1081	Db	 TTCTCTCAGCTCCAGCCCGGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCGAAGC 1140
1139	Qy	 AGTCCCGACCGTCCACCCCGACACCGCGGCTTCGCCCGCAGCGGGGCGGTATGGG 1198
1141	Db	 AGTCCCGACCGTCCACCCCGACACCGCGGCTTCGCCCGCAGCGGGGCGGTATGGG 1200
1199	Qy	 ACGAGGCATGCGCTTCGGCGCCCGGCTGCATCGCACCCCGGCGCGTGCCTGGACCCGCCGA 1258

1201	ACGAGGCATCTGCCTCTGGCGCCCGCTGCATCCGACCCGGCCCGCTGTCTGGACCCGCGCA	1266
1259	TGAAGCGGCTCCCAACGAGTGGCGGGCGGGCTTCCCCTCTTTCCATCTTCAAGCCCTCGC	1318
1261	TGAAGCGGCTCCCAACGAGTGGCGGGCGGGCTTCCCCTCTTTCCATCTTCAAGCCCTCGC	1320
1319	CGCCGATCCCGCCCGCCAGCCCGGCGGGCCGACCACTCTCGCTACGACCCGACCG	1378
1321	CGCCGATCCCGCCCGCCAGCCCGGCGGGCCGACCACTCTCGCTACGACCCGACCG	1380
1379	CCGCTCGCGCTCAGCTCGCTGGGAGCCGACCGCGCGGGGAGCCAGCCAGCGCGCG	1438
1381	CCGCTCGCGCTCAGCTCGCTGGGAGCCGACCGCGCGGGGAGCCAGCCAGCGCGCG	1440
1439	CGCTTGAGAGCACCCGTTACGGCTGCGCTGGCCAGAGGGGCGCCCGCTGGCCTTC	1498
1441	CGCTTGAGGGCCACCCGTACGGCTGCGCTGGCCAGAGGGGCGCCCGCTGGCCTTC	1500
1499	CGCCTCTGGCCTCAGCCCTCCCTACCGCTCCAGCCTGCTGGGCGAGAGTCCAGCC	1558
1501	CGCCTCTGGCCTCAGCCCTCCCTACCGCTCCAGCCTGCTGGGCGAGAGTCCAGCC	1560
1559	TGCGCTCGCGCCAGCAGAGCTCGTCTGCTGGCAGGGCACTGCTGCGGTGTCGGGG	1618
1561	TGCGCTCGCGCCAGCAGAGCTCGTCTGCTGGCAGGGCACTGCTGCGGTGTCGGGG	1620
1619	ACAAGCGGCTGCGACGACTACGGCTGGGACCTGCGAGGGCTGCAAGGGCTTTTCA	1678
1621	ACAAGCGGCTGCGACGACTACGGCTGGGACCTGCGAGGGCTGCAAGGGCTTTTCA	1680
1679	AGAAACAGTGCAGAAAAATGAAAATATGTTTGCCTGGCAATAAAAACTGCCAGTAG	1738
1681	AGAAACAGTGCAGAAAAATGAAAATATGTTTGCCTGGCAATAAAAACTGCCAGTAG	1740
1739	ACAAGAGCGTGGAAACGATGTCAGTCTCGATTTCAAGAGTGTCTCAGTGTGGAA	1798
1741	ACAAGAGCGTGGAAACGATGTCAGTCTCGATTTCAAGAGTGTCTCAGTGTGGAA	1800
1799	TGTTAAAGAAATGTTGTCGTACAGATAGTCTGAAAGGGAGGAGGTGCTCGCTTCCA	1858
1801	TGTTAAAGAAATGTTGTCGTACAGATAGTCTGAAAGGGAGGAGGTGCTCGCTTCCA	1860
1859	AACAAAGAGCCATTACACAGAAACCTTCTCAGCCCTCTCCACTTCTCCTCCCAATCT	1918
1861	AACAAAGAGCCATTACACAGAAACCTTCTCAGCCCTCTCCACTTCTCCTCCCAATCT	1920
1919	GCATGATGAATGCTTGTCTGGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTT	1978
1921	GCATGATGAATGCTTGTCTGGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTTT	1980
1979	CCAGATATGTTCCACTGACCGGCTGTGAGGCAACAGATGCTGAGCAGTGTGCAACAAT	2038
1981	CCAGATATGTTCCACTGACCGGCTGTGAGGCAACAGATGCTGAGCAGTGTGCAACAAT	2040
2039	TCTACAACTCTCGACGCTCCATTGATGTATCCAGAGCTGGGCGAGAAAAGATTCCGG	2098
2041	TCTACAACTCTCGACGCTCCATTTGATGTATCCAGAGCTGGGCGAGAAAGATTCCGG	2100
2099	GATTTACTGATCTCCCAAAGAGATCAGACATTTACTTTTGAATCAGCCTTTTGGAGC	2158
2101	GATTTACTGATCTCCCAAAGAGATCAGACATTTACTTTTGAATCAGCCTTTTGGAGC	2160
2159	TGTTTTGCTCAGACTTTCCATCAGGTCAACACTGCTGGAAGATAAGTTTGTGTTCTGCA	2218
2161	TGTTTTGCTCAGACTTTCCATCAGGTCAACACTGCTGGAAGATAAGTTTGTGTTCTGCA	2220
2219	ATGGAATTGTCTCGTCACTGACTTCACTGTCCTTCTGATTTGGGAGTGGCTCGACTCTA	2278
2221	ATGGAATTGTCTCGTCACTGACTTCACTGTCCTTCTGATTTGGGAGTGGCTCGACTCTA	2280
2279	TTAAAGACTTTTCCCTTAAATTTGACAGCCTGAACCTTGATATCCAGCCTTAGCCTGCC	2338
2281	TTAAAGACTTTTCCCTTAAATTTGACAGCCTGAACCTTGATATCCAGCCTTAGCCTGCC	2340

QY 2339 TGTGACACTGAGCATGATCAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAG 2398  
DB 2341 TGTGACACTGAGCATGATCAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAG 2400  
QY 2399 AGCTATGCAACAAGATCAACAAGCAGTTTAAAGAGACCACAGAGTAAAGGACAGGCTCTGG 2458  
DB 2401 AGCTATGCAACAAGATCAACAAGCAGTTTAAAGAGACCACAGAGTAAAGGACAGGCTCTGG 2460  
QY 2459 AGCCACCGAGTCCAAAGTCTCTGGTGCCCTGGTAGAACTGAGGAAGATCTGCACCCCTGG 2518  
DB 2461 AACCACCAAGTCCAAAGTCTCTGGTGCCCTGGTAGAACTGAGGAAGATCTGCACCCCTGG 2520  
QY 2519 GCTCCAGGCGATCTTCTACCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2578  
DB 2521 GCTCCAGGCGATCTTCTACCTGGAAGCTTGGTGTCTCCACCTTCCATCATTTG 2580  
QY 2579 ACAAGCTCTTCTGACACCCCTACCTTCTTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2638  
DB 2581 ACAAGCTCTTCTGACACCCCTACCTTCTTAATCAGAGCAGTGGAGCAGTGAAGTGCCT 2640  
QY 2639 CTTCTCTAGCAGCTGCTTGTGACGAGCAAAAGGATAGTTTGGAACTATCATTTCC 2698  
DB 2641 CCTCTCTAGCAGCTGCTTGTGACGAGCAAAAGGATAGTTTGGAACTATCATTTCC 2700  
QY 2699 TGTCTCTCTTAAAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2758  
DB 2701 TGTCTCTCTTAAAGGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2760  
QY 2759 GCTCTTTTCTTACAACTTAAAGCCAGAAATCTGCAGAGTATGTGTTGGGTTGTGT 2818  
DB 2761 GCTCTTTTCTTACAACTTAAAGCCAGAAATCTGCAGAGTATGTGTTGGGTTGTGT 2820  
QY 2819 TTATATTAGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA 2878  
DB 2821 TTATATTAGCATTTGGGGATGGGGTGGAGGGGTTATAGTTTATGAGGGTTTCTAA 2880  
QY 2879 GAAATGCTTAAAGACACTTTTGGACATGCTATCCAGCAGGAGGAAAAAGGATATA 2938  
DB 2881 GAAATGCTTAAAGACACTTTTGGACATGCTATCCAGCAGGAGGAAAAAGGATATA 2940  
QY 2939 TAACTGTTTTTAAACTCTTCTGGGAATCCAAATATAGTTCTGTTGATTTAAAAACA 2998  
DB 2941 TAACTGTTTTTAAACTCTTCTGGGAATCCAAATATAGTTCTGTTGATTTAAAAACA 3000  
QY 2999 GAACAGCCAAAGGTTGTTCCGACGGTAGGATGTGTTTAAAGATGGTCCCTGAAAT 3058  
DB 3001 GAACAGCCAAAGGTTGTTCCGACGGTAGGATGTGTTTAAAGATGGTCCCTGAAAT 3060  
QY 3059 ATGCTTCTGTATCAAGGTACGTATGTGGTGCACAAAGGAGGAGAACTTCTTTTAAAT 3118  
DB 3061 ATGCTTCTGTATCAAGGTACGTATGTGGTGCACAAAGGAGGAGAACTTCTTTTAAAT 3120  
QY 3119 TCCTTCTCTCTTTATTTTAAACAAATGGTGAAAGATGGAGGATTAACCTACAAATCAGACAT 3178  
DB 3121 TCCTTCTCTCTTTATTTTAAACAAATGGTGAAAGATGGAGGATTAACCTACAAATCAGACAT 3180  
QY 3179 GCAAAACAAATATGGCTGTTTCTTCCATAACCAAGTGCAATTTTTTAAAGTGTGTCT 3238  
DB 3181 GCAAAACAAATATGGCTGTTTCTTCCATAACCAAGTGCAATTTTTTAAAGTGTGTCT 3240  
QY 3239 TACTAAGTCTGTTTATTTAACTCTCTCTTTATTTCTATATGGAATATAAAGAGGAGGATCA 3298  
DB 3241 TACTAAGTCTGTTTATTTAACTCTCTCTTTATTTCTATATGGAATATAAAGAGGAGGATCA 3300  
QY 3299 TGTATGCAAAATCAGACGTTTAATATCCCTAGCAGAGCTGTGTTCACCTTCCCTGTGATC 3358  
DB 3301 TGTATGCAAAATCAGACGTTTAATATCCCTAGCAGAGCTGTGTTCACCTTCCCTGTGATC 3360  
QY 3359 CTTCTGAGGTATGGCCCATCAAGACTTTTGGCCATTTTATGATGGAAACAGATCCCTG 3418  
DB 3361 CTTCTGAGGTATGGCCCATCAAGACTTTTGGCCATTTTATGATGGAAACAGATCCCTG 3420

QY 3419 CCCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAACTGGATTACATGTAACTG 3478  
DB 3421 CCCTGACTGTCCAGCTATCTCTGAAAGTGGATCAGATTATAACTGGATTACATGTAACTG 3480  
QY 3479 TTTTGGTTGTGTTCTTATCAACCCACAGAGTTCCCTAAACTTGTCTCAGTTTATAGTAAC 3538  
DB 3481 TTTTGGTTGTGTTCTTATCAACCCACAGAGTTCCCTAAACTTGTCTCAGTTTATAGTAAC 3540  
QY 3539 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTCAGTTAGTATTTGATCCCTAGATAAG 3598  
DB 3541 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTCAGTTAGTATTTGATCCCTAGATAAG 3600  
QY 3599 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGTTAAGCACCAGGACAAATAGGATTTT 3658  
DB 3601 AACATGCAAAATCAGCAGGAACTGGTTCATACAGGTTAAGCACCAGGACAAATAGGATTTT 3660  
QY 3659 TATAGATATAATTTAAATTTTGTGTTA- - TTGGTTAAGGAGA- -CAATTTTGGAGAGCAAGCA 3715  
DB 3661 TATAGATATAATTTAAATTTTGGTAAATTTGGTTAAGGAGACCAATTTTGGAGAGCAAGCA 3720  
QY 3716 AA- - -TCTTTTAAATAATAGTATGAATGTGAATACTAGAAAGATTTTAAATAATAGTAT 3772  
DB 3721 AATCTTCTTTTAAATAATAGTATGAATGTGAATACTAGAAAGATTTTAAATAATAGTAT 3780  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
DB 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 3

US-09-853-386-78  
; Sequence 78, Application US/09853386  
; Patent No. US20020049151A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Evelyn  
; APPLICANT: Bresnihan, Barry  
; APPLICANT: Conneely, Orla  
; APPLICANT: Fitzgerald, Oliver  
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR  
; TITLE OF INVENTION: Subfamily of Nuclear Transcription Factors  
; FILE REFERENCE: P01972US1  
; CURRENT APPLICATION NUMBER: US/09/853,386  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/203645  
; PRIOR FILING DATE: 2000-05-12  
; NUMBER OF SEQ ID NOS: 153  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 78  
; LENGTH: 3802  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-09-853-386-78

Query Match 98.1%; Score 3722; DB 9; Length 3802;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 1 ATAATGAGTGCCGAGAGAGCGGAAACGCGAGCGGGAGAGCGGAGTCTCTGCT 60  
DB 1 ATAAATGAGTGCCGAGAGAGCGGAAACGCGAGCGGGAGAGCGGAGTCTCTGCT 60  
QY 61 CCGGCCCCCACCCTCCAGCTCTCTGCTCTCTCCGCTCCCATACACAGCGGCTCA 120  
DB 61 CCGGCCCCCACCCTCCAGCTCTCTGCTCTCTCCGCTCCCATACACAGCGGCTCA 120  
QY 121 CACCGCT 178  
DB 121 CACCGCT 180  
QY 179 ACTTGGCT 238  
DB 181 ACTTGGCT 240



Db 2401 AGCTATGCAACAGATCAACAGCAGTTTAAAGACCACCCAGAGTAAGGGACAGGCTCTGG 2460  
Qy 2459 AGCCCAACGAGTCCAAAGGTCCTGGGTGCCCTGGTAGAACTGAGGAAGATCTGCACCCCTGG 2518  
Db 2461 AACCCAAACGAGTCCAAAGGTCCTGGGTGCCCTGGTAGAACTGAGGAAGATCTGCACCCCTGG 2520  
Qy 2519 GCTCCAGCGCATCTTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATG 2578  
Db 2521 GCTCCAGCGCATCTTCTACCTGAAGCTGGAAGCTTGGTGTCTCCACCTTCCATCATG 2580  
Qy 2579 ACAAGCTCTTCTCGGACACCCCTACCTTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2638  
Db 2581 ACAAGCTCTTCTCGGACACCCCTACCTTCTAATCAGGAGCAGTGGAGCAGTGAAGTGCCT 2640  
Qy 2639 CCTCTCCTAGCAGCTCTTCTAGCAGCAAAAGGATAGGTTTGGAAAACCTATCATTTCC 2698  
Db 2641 CCTCTCCTAGCAGCTCTTCTAGCAGCAAAAGGATAGGTTTGGAAAACCTATCATTTCC 2700  
Qy 2699 TGTCTCTCTTAAGAGGAAAGCAGCTCTGTAGAAGCAAGACATCTCTTTTTCG 2758  
Db 2701 TGTCTCTCTTAAGAGGAAAGCAGCTCTGTAGAAGCAAGACATCTCTTTTTCG 2760  
Qy 2759 GCTCTTTTCTTACAACTTAAAGCCAGAAAACCTTGCAAGATATTGTGTGGGTGTGTT 2818  
Db 2761 GCTCTTTTCTTACAACTTAAAGCCAGAAAACCTTGCAAGATATTGTGTGGGTGTGTT 2820  
Qy 2819 TTATATTAGGCAATTGGGGATGGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2878  
Db 2821 TTATATTAGGCAATTGGGGATGGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2880  
Qy 2879 GAAATGCTAACAAAGCACTTTGGCAATGCTATCCAGCAGGAGGAAAAAGGATAATA 2938  
Db 2881 GAAATGCTAACAAAGCACTTTGGCAATGCTATCCAGCAGGAGGAAAAAGGATAATA 2940  
Qy 2939 TAACTGTTTTAAACCTTTCTGGGAAATCCAATTATAGTGTCTTGTATTTAAAAACAA 2998  
Db 2941 TAACTGTTTTAAACCTTTCTGGGAAATCCAATTATAGTGTCTTGTATTTAAAAACAA 3000  
Qy 2999 GAACAGCCAAAGGTTTGGCCAGGGTAGATGTGCTTAAAGATGTCCTTGAAT 3058  
Db 3001 GAACAGCCAAAGGTTTGGCCAGGGTAGATGTGCTTAAAGATGTCCTTGAAT 3060  
Qy 3059 ATGCTTCCTGATCAAGAGTAGTATGTGTCGCAACAGGAGGAGAACTTCTTTAAT 3118  
Db 3061 ATGCTTCCTGATCAAGAGTAGTATGTGTCGCAACAGGAGGAGAACTTCTTTAAT 3120  
Qy 3119 TCCTTCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATACCTACAAATCAGACAT 3178  
Db 3121 TCCTTCTCTTTATTTTAAACAAATGGTGAAGATGGAGGATACCTACAAATCAGACAT 3180  
Qy 3179 GGCAAAACAAATATGCTGTTTCCATAAACAAGTGCATTTTAAAGTGTGCT 3238  
Db 3181 GGCAAAACAAATATGCTGTTTCCATAAACAAGTGCATTTTAAAGTGTGCT 3240  
Qy 3239 TACTAGTCTCTTTTAACTCTCTTTTATCTATATGGAATATAAGAGGAGGAGTCA 3298  
Db 3241 TACTAGTCTCTTTTAACTCTCTTTTATCTATATGGAATATAAGAGGAGGAGTCA 3300  
Qy 3299 TGTAGCAATGACACGTTTAAATCCCTAGCAGAGCTGTGTTACCTTCCCTGTGATC 3358  
Db 3301 TGTAGCAATGACACGTTTAAATCCCTAGCAGAGCTGTGTTACCTTCCCTGTGATC 3360  
Qy 3359 CCTTCTGAGGTATGGCCCATCAAGACTTTTAAAGCCATTTCTTGTAGGAAACAGATCCCTG 3418  
Db 3361 CCTTCTGAGGTATGGCCCATCAAGACTTTTAAAGCCATTTCTTGTAGGAAACAGATCCCTG 3420  
Qy 3419 CCTGACTGTCAGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAC 3478  
Db 3421 CCTGACTGTCAGCTATCCTGAAAGTGGATCAGATTATAAAGTGGATTACATGTAAC 3480  
Qy 3479 TTTTGGTGTCTTATCAACCCACAGTTCCTTAACTTGTCTTCAAGTTATAGTAAC 3538  
Db 3481 TTTTGGTGTCTTATCAACCCACAGTTCCTTAACTTGTCTTCAAGTTATAGTAAC 3540

Qy 3539 TGACTGGTATATATTCATTTCAGAAAGCCCATAAAGTCAAGTTCAGTATTTTGTATCCCTAGATAAG 3598  
Db 3541 TGACTGGTATATATTCATTTCAGAAAGCCCATAAAGTCAAGTTCAGTATTTTGTATCCCTAGATAAG 3600  
Qy 3599 AACATGCAAAATCAGCAGGAACCTGGTTCATACAGGGTAAGCACCAGGACAATAAGGATTTT 3658  
Db 3601 AACATGCAAAATCAGCAGGAACCTGGTTCATACAGGGTAAGCACCAGGACAATAAGGATTTT 3660  
Qy 3659 TATAGATATAATTTAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3715  
Db 3661 TATAGATATAATTTAATTTTGTTA--TTGGTTAAGGAGA--CAATTTTGGAGAGCAAGCA 3720  
Qy 3715 AA---TCITTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTTAAAAAATAGTAT 3772  
Db 3721 AATCTTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAAGATTTTAAAAAATAGTAT 3780  
Qy 3773 GAGTGTGAGTACTAGGNAAGAT 3794  
Db 3781 GAGTGTGAGTACTAGGNAAGAT 3802

## RESULT 4

US-10-414-080-1  
; Sequence 1, Application US/10414080  
; Publication No. US20030220288A1  
; GENERAL INFORMATION:  
; APPLICANT: MULLICAN, SHANNON E.  
; APPLICANT: CONNELLY, ORLA M.  
; APPLICANT: MILBRANDT, JEFFREY  
; TITLE OF INVENTION: NOR-1 AND NUR77 NUCLEAR RECEPTORS AS TARGETS FOR  
; FILE REFERENCE: P02454US1  
; CURRENT APPLICATION NUMBER: US/10/414,080  
; PRIOR FILING DATE: 2003-04-15  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3802  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-414-080-1

Query Match 98.1%; Score 3722; DB 17; Length 3802;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

Qy 1 ATAAATGACGTCCGAGAGAGCGGAAACGGCAGCCGGAGAGCGGAGTCTCTGCCT 60  
Db 1 ATAAATGACGTCCGAGAGAGCGGAAACGGCAGCCGGAGAGCGGAGTCTCTGCCT 60  
Qy 61 CCGGCCCCCACCCTCCAGCTCCTCTCTCCGCTCCCATACACAGACGGCTCA 120  
Db 61 CCGGCCCCCACCCTCCAGCTCCTCTCTCCGCTCCCATACACAGACGGCTCA 120  
Qy 121 CACCGCTCCTCTCACTCGCACACAGACACAAAGCGCGCACACAGGCTCCGACACAC 178  
Db 121 CACCGCTCCTCTCACTCGAACACACAGACACAAAGCGCGCACACAGGCTCCGACACAC 180  
Qy 179 ACTTCGCTCTCCCGCGCTCACACCTCTTTGCGCTGAGCCCTTGGCGGTGAGCGCG 238  
Db 181 ACTTCGCTCTCCCGCGCTCACACCTCTTTGCGCTGAGCCCTTGGCGGTGAGCGCG 240  
Qy 239 GCGCGCAGCTGAGCCCTCCCGGCTCACTTTGCAAGCTGACGGTCCGCGAGTGGC 298  
Db 241 GCGCGCAGCTGAGCCCTCCCGGCTCACTTTGCAAGCTGACGGTCCGCGAGTGGC 300  
Qy 299 CGTGAGGTGGAAACAGCGGGCGCATCTCTCCCTTGTTCACAGCCCAAGCAGGACGCC 358  
Db 301 CGTGAGGTGGAAACAGCGGGCGCATCTCTCCCTTGTTCACAGCCCAAGCAGGACGCC 360

QY 359 CGCGGAACTCTCGGTGCTCTCCCATGATCGGGATCGCAGCATCCCCACGACCG 418  
DB |||||  
QY 361 CGCGGAACTCTCGGTGCTCTCCCATGATCGGGATCGCAGCATCCCCACGACCG 420  
DB |||||  
QY 419 CTCACGCGCTCGCGGAGCGCTGGGCTTTGTATACCGCAGCCCTTCGGGACAGCAGCTGT 478  
DB |||||  
QY 421 CTCACGCGCTCGCGGAGCGCTGGGCTTTGTATACCGCAGCCCTTCGGGACAGCAGCTGT 480  
DB |||||  
QY 479 GACTCCCGCCAGTGAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGGAAC 538  
DB |||||  
QY 481 GACTCCCGCCAGTGAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGGAAC 540  
DB |||||  
QY 539 CGCGCAGCACTCAAGCCCACTCGGGAAGAGCGAGCCCGGAGCCCGGCGCTGAGC 598  
DB |||||  
QY 541 CGCGCAGCACTCAAGCCCACTCGGGAAGAGCGAGCCCGGAGCCCGGCGCTGAGC 600  
DB |||||  
QY 599 CTGGACCTTTAGCGGTGCGCGGAGCACTGCGCGGCGCTTCGGCTCGCGGAGCTCGGCTC 658  
DB |||||  
QY 601 CTGGACCTTTAGCGGTGCGCGGAGCACTGCGCGGCGCTTCGGCTCGCGGAGCTCGGCTC 660  
DB |||||  
QY 659 CTCCTACACTCTCAGCGCTCGGTGAGAGACCCCGGAGCCCGGAGCTCAGCGCGCAAGAT 718  
DB |||||  
QY 661 CTCCTACACTCTCAGCGCTCGGTGAGAGACCCCGGAGCCCGGAGCTCAGCGCGCAAGAT 720  
DB |||||  
QY 719 ACCCTCCAGATATGCCCTCGCTCCAGCCCAATATAGCCCTTCGGCTCGCGGAGCTTCAGTT 778  
DB |||||  
QY 721 ACCCTCCAGATATGCCCTCGCTCCAGCCCAATATAGCCCTTCGGCTCGCGGAGCTTCAGTT 780  
DB |||||  
QY 779 ATGCGCGGAGACATACAGCTCGGAATACACCGGAGATCATGAACCCCGGACCTACACA 838  
DB |||||  
QY 781 ATGCGCGGAGACATACAGCTCGGAATACACCGGAGATCATGAACCCCGGACCTACACA 840  
DB |||||  
QY 839 AGCTGACCATGAGCTTGGGAGGAGTACGATCAGCGCTACGCGACCACTCCCTGCGCCA 898  
DB |||||  
QY 841 AGCTGACCATGAGCTTGGGAGGAGTACGATCAGCGCTACGCGACCACTCCCTGCGCCA 900  
DB |||||  
QY 899 GCATCAGTACTCTGCTGGAGGGTACTCGAGCAACTACGAACTCAAGCCCTTCGCGTGT 958  
DB |||||  
QY 901 GCATCAGTACTCTGCTGGAGGGTACTCGAGCAACTACGAACTCAAGCCCTTCGCGTGT 960  
DB |||||  
QY 959 ACCAAATGACGCGGCTTTGATCAAAAGTGAGAGGGGGGGGGGGGGGGGGGGGGGGGG 1018  
DB |||||  
QY 961 ACCAAATGACGCGGCTTTGATCAAAAGTGAGAGGGGGGGGGGGGGGGGGGGGGGGGG 1020  
DB |||||  
QY 1019 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1078  
DB |||||  
QY 1021 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1080  
DB |||||  
QY 1079 TTCTTCAGGCTCCAGCGCGGAGGAGGAGTGTGCTGCCAGCACCTCCATGTACTTCAAGC 1138  
DB |||||  
QY 1081 TTCTTCAGGCTCCAGCGCGGAGGAGGAGTGTGCTGCCAGCACCTCCATGTACTTCAAGC 1140  
DB |||||  
QY 1139 AGTCCCAACCGTCCACCCCAACCGCGGCTTCCTCCCGGAGGGGGGGGGGGGGGGGGGG 1198  
DB |||||  
QY 1141 AGTCCCAACCGTCCACCCCAACCGCGGCTTCCTCCCGGAGGGGGGGGGGGGGGGGGGG 1200  
DB |||||  
QY 1199 ACGAGGCACTGCTCGGCGCGCGGCTGCTGCAACCGCGCGCGCTGCTGCAACCGCGCGG 1258  
DB |||||  
QY 1201 ACGAGGCACTGCTCGGCGCGCGGCTGCTGCAACCGCGCGCGCTGCTGCAACCGCGCGG 1260  
DB |||||  
QY 1259 TGAAGGCGGTCCCAACCGGTGCGGCGGCTTCGGCTTCCTCCATCTTCAAGCCCTGCG 1318  
DB |||||  
QY 1261 TGAAGGCGGTCCCAACCGGTGCGGCGGCTTCGGCTTCCTCCATCTTCAAGCCCTGCG 1320  
DB |||||  
QY 1319 GCGCGCATCCCG 1378  
DB |||||  
QY 1321 GCGCGCATCCCG 1380  
DB |||||  
QY 1379 CCGCTGCGGCGCTCAGCGCTGCGGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1438  
DB |||||  
QY 1381 CCGCTGCGGCGCTCAGCGCTGCGGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440  
DB |||||  
QY 1439 CGCTTGAGAGCCACCGGTACGCGGTGCGGTGCGGCAAGAGGGGGGGGGGGGGGGGGGG 1498  
DB |||||

DB 1441 CGCTTGAGGGCCACCCGTAACGGGCTGCGCTGCGCAAGAGGGGGGGGGGGGGGGGGGG 1500  
QY |||||  
QY 1499 CGCTCTCGGCTTACCGCTCCCTTACCGGCTCCAGCTGCTGGGGAGAGTCCAGGCC 1558  
DB |||||  
QY 1501 CGCTCTCGGCTTACCGCTCCCTTACCGGCTCCAGCTGCTGGGGAGAGTCCAGGCC 1560  
DB |||||  
QY 1559 TGCGCTCGCGCCCGCAGCAGGAGCTCGTCTGCTGGCGAGGGCACGTGTGCCGTGTGCCGG 1618  
DB |||||  
QY 1561 TGCGCTCGCGCCCGCAGCAGGAGCTCGTCTGCTGGCGAGGGCACGTGTGCCGTGTGCCGG 1620  
DB |||||  
QY 1619 ACAACGCGCTGCGCAGCATTAACGCGCTGCGAACCTCGGAGGGCTGCAAGGGCTTTTCA 1678  
DB |||||  
QY 1621 ACAACGCGCTGCGCAGCATTAACGCGCTGCGAACCTCGGAGGGCTGCAAGGGCTTTTCA 1680  
DB |||||  
QY 1679 AGAGAACAGTGCAGAAATATGTTTCCCTGGCAATAAATAAATGCCCCCAGTAG 1738  
DB |||||  
QY 1681 AGAGAACAGTGCAGAAATATGTTTCCCTGGCAATAAATAAATGCCCCCAGTAG 1740  
DB |||||  
QY 1739 ACAAGAGACGTCGAAACCGATGTCTAGTACTGTCTGATTTTCAAGAGTGTCTAGTGTGAA 1798  
DB |||||  
QY 1741 ACAAGAGACGTCGAAACCGATGTCTAGTACTGTCTGATTTTCAAGAGTGTCTAGTGTGAA 1800  
DB |||||  
QY 1799 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGGGAGAGAGTGTCTGCTTCCA 1858  
DB |||||  
QY 1801 TGGTAAAGAAAGTTGTCCGTACAGATAGTCTGAAGGGAGAGAGTGTCTGCTTCCA 1860  
DB |||||  
QY 1859 AACCAAGAGCGCCATTACAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCCAATCT 1918  
DB |||||  
QY 1861 AACCAAGAGCGCCATTACAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCTCCAATCT 1920  
DB |||||  
QY 1919 GCATGATGAATGCCCCCTGTCTCGAGCTTAAACAGACTCAACACCCAGAGATCTTGAATTT 1978  
DB |||||  
QY 1921 GCATGATGAATGCCCCCTGTCTCGAGCTTAAACAGACTCAACACCCAGAGATCTTGAATTT 1980  
DB |||||  
QY 1979 CCAGATCTCTCCACTGACCCAGGCTGCTGAGGACAGAGTGTGAGCATGTGCAACAT 2038  
DB |||||  
QY 1981 CCAGATCTCTCCACTGACCCAGGCTGCTGAGGACAGAGTGTGAGCATGTGCAACAT 2040  
DB |||||  
QY 2039 TCTCAACCTCTGACAGCTCCATTTGATGTATCCAGAGCTGGGACAGAAAGATTCGG 2098  
DB |||||  
QY 2041 TCTCAACCTCTGACAGCTCCATTTGATGTATCCAGAGCTGGGACAGAAAGATTCGG 2100  
DB |||||  
QY 2099 GATTTACTGATCTCCCAAGAGATCAGATTAATTTGAATCAGCCCTTTTGGAGC 2158  
DB |||||  
QY 2101 GATTTACTGATCTCCCAAGAGATCAGATTAATTTGAATCAGCCCTTTTGGAGC 2160  
DB |||||  
QY 2159 TGTTTGTCTCAGACTTTCCATCAGTCAACACCTGCTGAAGATAAGTTTGTGTTCTGCA 2218  
DB |||||  
QY 2161 TGTTTGTCTCAGACTTTCCATCAGTCAACACCTGCTGAAGATAAGTTTGTGTTCTGCA 2220  
DB |||||  
QY 2219 ATGGACTTGTCTCAGATCGACTTTCAGTGCCTTTCGTGGAATTTGGGAGTGGCTCGACTTA 2278  
DB |||||  
QY 2221 ATGGACTTGTCTCAGATCGACTTTCAGTGCCTTTCGTGGAATTTGGGAGTGGCTCGACTTA 2280  
DB |||||  
QY 2279 TTAAGACTTTTCTTAAATTTGAGAGCTGAACTTTGATATCAAGCCCTTAGCCTGC 2338  
DB |||||  
QY 2281 TTAAGACTTTTCTTAAATTTGAGAGCTGAACTTTGATATCAAGCCCTTAGCCTGC 2340  
DB |||||  
QY 2339 TGTGAGCTGAGCATGATCAGAAAGACATGGGTAAAGAAACCAAGAGAGTCCGAAG 2398  
DB |||||  
QY 2341 TGTGAGCTGAGCATGATCAGAAAGACATGGGTAAAGAAACCAAGAGAGTCCGAAG 2400  
DB |||||  
QY 2399 AGCTATGCAACAGATCAACAGCTTTAAAGACCAACAGATTAAGGAGAGGCTCTGG 2458  
DB |||||  
QY 2401 AGCTATGCAACAGATCAACAGCTTTAAAGACCAACAGATTAAGGAGAGGCTCTGG 2460  
DB |||||  
QY 2459 AGCCCAACGAGTCAAGGCTTCCCTGGTCCCTGCTGAGACTGAGGAGATCTGCACCCCTCG 2518  
DB |||||  
QY 2461 AAGCCCAACGAGTCAAGGCTTCCCTGGTCCCTGCTGAGACTGAGGAGATCTGCACCCCTCG 2520  
DB |||||  
QY 2519 GCCTCCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCTTG 2578  
DB |||||

Db 2521 GCCTCAGCGCATCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2580  
QY 2579 ACAAGCTCTCTCGACACCCCTACCTTTCTAATCAGGACGAGTGGAGCAGTGCCT 2638  
Db 2581 ACAAGCTCTCTCGACACCCCTACCTTTCTAATCAGGACGAGTGGAGCAGTGCCT 2640  
QY 2639 CCTCTCCTAGCACCTCTGCTTACCGACGAAAGGATAGGTTTGGAAACCTATCATTTCC 2698  
Db 2641 CCTCTCCTAGCACCTCTGCTTACCGACGAAAGGATAGGTTTGGAAACCTATCATTTCC 2700  
QY 2699 TGTCTCTTCCCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAAAGACTTCTTTTTCCTG 2758  
Db 2701 TGTCTCTTCCCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAAAGACTTCTTTTTCCTG 2760  
QY 2759 GCTCTTTTCCCTTACAACTTAAAGCCAGAAACCTTCGACAGTATGTGTTGGGGTGTGTT 2818  
Db 2761 GCTCTTTTCCCTTACAACTTAAAGCCAGAAACCTTCGACAGTATGTGTTGGGGTGTGTT 2820  
QY 2819 TTATATTTAGGCATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTTCCTAA 2878  
Db 2821 TTATATTTAGGCATTTGGGGATGGGGTGGAGGGGTATAGTTTATGAGGGTTCCTAA 2880  
QY 2879 GAAATTTGCTTACAAAGCACTTTTGGCAATGCTATCCACGACGAGGAAAAAGGATAATA 2938  
Db 2881 GAAATTTGCTTACAAAGCACTTTTGGCAATGCTATCCACGACGAGGAAAAAGGATAATA 2940  
QY 2939 TAACTGTTTTAAACCTCTTTTGGGGAATCCAAATATAGTTGCTTTGTTATTTAAACAA 2998  
Db 2941 TAACTGTTTTAAACCTCTTTTGGGGAATCCAAATATAGTTGCTTTGTTATTTAAACAA 3000  
QY 2999 GAACAGCCAAAGGTTGTTCCGACAGGTAGGTGCTTTAAAGATGGTCCCTTCAAAAT 3058  
Db 3001 GAACAGCCAAAGGTTGTTCCGACAGGTAGGTGCTTTAAAGATGGTCCCTTCAAAAT 3060  
QY 3059 ATGCTTCTCTGATCAAAAGTACGTATGTGCTGCAAAAGGACAGAACTTCTTTTAAAT 3118  
Db 3061 ATGCTTCTCTGATCAAAAGTACGTATGTGCTGCAAAAGGACAGAACTTCTTTTAAAT 3120  
QY 3119 TCCTTCTCTCTTTATTTAAACAAATGGTGAAGATGGAGATTAAGTACAAATCAGACAT 3178  
Db 3121 TCCTTCTCTCTTTATTTAAACAAATGGTGAAGATGGAGATTAAGTACAAATCAGACAT 3180  
QY 3179 GCGAAACAAATATGGCTGTTGCTTCCATAACAGTGCATTTTAAAGTGTGCT 3238  
Db 3181 GCGAAACAAATATGGCTGTTGCTTCCATAACAGTGCATTTTAAAGTGTGCT 3240  
QY 3239 TACTAAGTCTGTTTATTAACCTCTCCTTTTATCTATATGAAATAAAGAGGAGTCA 3298  
Db 3241 TACTAAGTCTGTTTATTAACCTCTCCTTTTATCTATATGAAATAAAGAGGAGTCA 3300  
QY 3299 TGTTAGCAATGACACGTTTATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3358  
Db 3301 TGTTAGCAATGACACGTTTATATCCCTAGCAGAGCTGTGTTCACTTCCCTGTGATC 3360  
QY 3359 CCTTCTGAGTATGCGCCATCAAGACTTTTAGGCCATTTCTGTATGGAACCAAGATCCCTG 3418  
Db 3361 CCTTCTGAGTATGCGCCATCAAGACTTTTAGGCCATTTCTGTATGGAACCAAGATCCCTG 3420  
QY 3419 CCTGACTCTCCAGCTATCCTGAAAGTGGATCAGATTATATAACTGGATTACATGTAAC 3478  
Db 3421 CCTGACTCTCCAGCTATCCTGAAAGTGGATCAGATTATATAACTGGATTACATGTAAC 3480  
QY 3479 TTTTGGTGTGTTTATCAACCCCAACAGAGTTCCCTAACTTGTCTTCAAGTTATAGTAAC 3538  
Db 3481 TTTTGGTGTGTTTATCAACCCCAACAGAGTTCCCTAACTTGTCTTCAAGTTATAGTAAC 3540  
QY 3539 TGACTGGTATATTCATTAGAGCGCCATTAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3598  
Db 3541 TGACTGGTATATTCATTAGAGCGCCATTAAGTCAAGTTCAGTATTTGATCCCTAGATAAG 3600  
QY 3599 AACATGCAATCAGCAGGAACTGGTTCATACAGGGTAAGCACAGGAGCAATAAGGATTTT 3658  
Db 3601 AACATGCAATCAGCAGGAACTGGTTCATACAGGGTAAGCACAGGAGCAATAAGGATTTT 3660

QY 3659 TATAGATATATATTTAATTTTGTGTTA--TTGGTTAAGGAGA-CAATTTTGGAGAGCAAGCA 3715  
Db 3661 TATAGATATATATTTAATTTTGTGTTAATTTGGTAAATTTGGGTTAAGGAGACCAATTTTGGAGAGCAAGCA 3720  
QY 3716 AA---TCCTTTTAAAAAATAGTATGAATCTGTAATACTAGAAAAAGATTTAAAAAATAGTAT 3772  
Db 3721 AATCTTCTTTTAAAAAATAGTATGAATCTGTAATACTAGAAAAAGATTTTAAAGAAATAGTAT 3780  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

RESULT 5

US-10-659-004-115  
; Sequence 115, Application US/10659004  
; Publication No. US20050048507A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhong et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-608  
; CURRENT APPLICATION NUMBER: US/10/659,004  
; CURRENT FILING DATE: 2003-09-09  
; PRIOR APPLICATION NUMBER: 60/295,607  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/295,661  
; PRIOR FILING DATE: 2001-06-04  
; PRIOR APPLICATION NUMBER: 60/296,404  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/296,418  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: 60/297,414  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: 60/297,567  
; PRIOR FILING DATE: 2001-06-12  
; PRIOR APPLICATION NUMBER: 60/298,285  
; PRIOR FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: 60/298,556  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/299,949  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: 60/300,883  
; PRIOR FILING DATE: 2001-06-26  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 187  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 115  
; LENGTH: 3802  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (732)..(2609)  
US-10-659-004-115

Query Match 98.1%; Score 3722; DB 19; Length 3802;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3779; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 1 ATAAATGAGTGTCCGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGTCTCCTGCT 60  
Db 1 ATAAATGAGTGTCCGAGAGGAGCGGAGCGGAGCGGAGCGGAGTCTCCTGCT 60  
QY 61 CCGCGCCCCCACCCTCCAGCTCCCTGCTCCTCCTCGCTCCCATACACAGACGCGCTCA 120  
Db 61 CCGCGCCCCCACCCTCCAGCTCCCTGCTCCTCCTCGCTCCCATACACAGACGCGCTCA 120  
QY 121 CACCGCTCCTCTCCTCCTCGCACACAGACAAAGCGCGCACACAGGCTCCG--CACACAC 178  
Db 121 CACCGCTCCTCTCCTCCTCGAACACACAGACAAAGCGCGCACACAGGCTCCGACACAC 180  
QY 179 ACTTGTCTCTCCCGCGGCTCACACCCCTCTTGTGCTGAGCCCTTGTCCGCTGAGCGCGG 238



181	ACTTCGCTCTCCGCGCGCTCACACCTCTTGGCCCTGAGCCCTTGCCTGTCAGCGCG	240	1261	TGAAGCGGCTCCCTCCACGCTGGCGCGCGCTTCCCGCTCTTCCACCTTCAAGCCCTCGC	1320
239	CGCCGAGCTGAGCGCCCTCCCGGCTCACTTTGCAACGCTGACGCTGCGCGAGTGGC	298	1319	CGCGCATCCCG	1378
241	CGCCGAGCTGAGCGCCCTCCCGGCTCACTTTGCAACGCTGACGCTGCGCGAGTGGC	300	1321	CGCGCATCCCG	1380
299	CGTGGAGTGGGAAACAGCGGGGCACTCTCCCGCTTGTCTACAGCCCAAGCCAGGACGC	358	1379	CGCTCGCGGCTCAGCTTCCCGCTTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1438
301	CGTGGAGTGGGAAACAGCGGGGCACTCTCCCGCTTGTCTACAGCCCAAGCCAGGACGC	360	1381	CGCTCGCGGCTCAGCTTCCCGCTTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG	1440
359	CGCGAACTCTCGGCTGTCTCCATGAGTGGGATCGAGCATCCCGCCAGCGCG	418	1439	CGTTGAGAGCAACCGTACCGGCTCGCGTGGCGCAAGAGGGCGCGCGCGCGCTTCC	1498
361	CGCGAACTCTCGGCTGTCTCCATGAGTGGGATCGAGCATCCCGCCAGCGCG	420	1441	CGTTGAGAGCAACCGTACCGGCTCGCGTGGCGCAAGAGGGCGCGCGCGCTTCC	1500
419	CTCACCGCTCCGGAGCGCTGGCTTTGTACACCGCAGCCCTTCCGGGACAGCAGTGT	478	1499	CGCTCTCGGCTCAGCGCTTCCCGCTTACCGCGTTCAGCGCTTGGCGGAGAGTCCAGCG	1558
421	CTCACCGCTCCGGAGCGCTGGCTTTGTACACCGCAGCCCTTCCGGGACAGCAGTGT	480	1501	CGCTCTCGGCTCAGCGCTTCCCGCTTACCGCGTTCAGCGCTTGGCGGAGAGTCCAGCG	1560
479	GACTCCCCCAGTGAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGAAC	538	1559	TGCGCTCGCGCGCAGCAGGAGCTCGTGTCTGGCGAGGCGCAGTGTGCGGTGGGG	1618
481	GACTCCCCCAGTGAGATTTGGGACAGCTCTCTAGAACTCGCTCTAAGAGCGAAC	540	1561	TGCGCTCGCGCGCAGCAGGAGCTCGTGTCTGGCGAGGCGCAGTGTGCGGTGGGG	1620
539	CGCCACAGCACTCAAGCCCACTGGCGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC	598	1619	ACAGCGCGCTTCCCGCAGCACTACCGGCTGCGAACCTTCCGAGGCGTGCAGGGCTTTTCA	1678
541	CGCCACAGCACTCAAGCCCACTGGCGAAGAGGCGAGCCCGGCAAGCCCGGCGCTGAGC	600	1621	ACAGCGCGCTTCCCGCAGCACTACCGGCTGCGAACCTTCCGAGGCGTGCAGGGCTTTTCA	1680
599	CTGGACCTTAGCGGTGCGGGGAGCAGCTGCGGGGCTTGGCTCGCGGAGCTCGGCTC	658	1679	AGAGAACAGTGCAGAAATATGTTTTGCTTGGCAATAAATAAAGTGCAGTAG	1738
601	CTGGACCTTAGCGGTGCGGGGAGCAGCTGCGGGGCTTGGCTCGCGGAGCTCGGCTC	660	1681	AGAGAACAGTGCAGAAATATGTTTTGCTTGGCAATAAATAAAGTGCAGTAG	1740
659	CTCCTACACTCTCAGCTCGCTGGAGAGACCCCGAGCCCACTTACAGCGCGCAAGAT	718	1739	ACAAGAGCGTGCAGAACCGATGTCTAGTCTCGATTTTCAAGAGTGTCTCAGTGTGGAA	1798
661	CTCCTACACTCTCAGCTCGCTGGAGAGACCCCGAGCCCACTTACAGCGCGCAAGAT	720	1741	ACAAGAGCGTGCAGAACCGATGTCTAGTCTCGATTTTCAAGAGTGTCTCAGTGTGGAA	1800
719	ACCTCTCAGATATGCGCTGCTCCAGCGCAATATAGCCCTTCCCTCCAGTTCAGTT	778	1799	TGTTAAAGAGTGTCTCGTACAGATAGTCTGAAAGGGAGGAGAGTGTCTGCTTCCA	1858
721	ACCTCTCAGATATGCGCTGCTCCAGCGCAATATAGCCCTTCCCTCCAGTTCAGTT	780	1801	TGTTAAAGAGTGTCTCGTACAGATAGTCTGAAAGGGAGGAGAGTGTCTGCTTCCA	1860
779	ATCGGGGAGACATACAGCTCGGATACACGCGAGATCATGAACCCCGACTACCA	838	1859	AACCAAGAGCGCTTCAACAGGAACTTCTCAGCGCTTCCACCTTCTCCTCAATCT	1918
781	ATCGGGGAGACATACAGCTCGGATACACGCGAGATCATGAACCCCGACTACCA	840	1861	AACCAAGAGCGCTTCAACAGGAACTTCTCAGCGCTTCCACCTTCTCCTCAATCT	1920
839	AGTGCACATGAGCTTGGCAGCACTGAGATACCGGCTACAGCACACGCTCCCTGCCA	898	1919	GCATGATGAATGCTTGTTCGAGCTTTAAGAGTCTCAACACCCAGAGATCTTGAATTT	1978
841	AGTGCACATGAGCTTGGCAGCACTGAGATACCGGCTACAGCACACGCTCCCTGCCA	900	1921	GCATGATGAATGCTTGTTCGAGCTTTAAGAGTCTCAACACCCAGAGATCTTGAATTT	1980
899	GCATCAGTACTTGTGGAGGGCTACTCGAGCACTACGAACTCAAGCCCTTCTGCGTGT	958	1979	CCAGATCTGTCCCACTGACCGAGCTGTGCGAGGCGAGATGTGAGCATGTGCAACAAT	2038
901	GCATCAGTACTTGTGGAGGGCTACTCGAGCACTACGAACTCAAGCCCTTCTGCGTGT	960	1981	CCAGATCTGTCCCACTGACCGAGCTGTGCGAGGCGAGATGTGAGCATGTGCAACAAT	2040
959	ACCAAAATGACGGCCCTTCATCAAGTGGAGGGGGCGCGCGCGCGCTACCATCAC	1018	2039	TCTACAACTCTGACAGCCCTCCATTTGATGTATCCAGAGCTGGGCGAGAAAGATTCGG	2098
961	ACCAAAATGACGGCCCTTCATCAAGTGGAGGGGGCGCGCGCGCGCTACCATCAC	1020	2041	TCTACAACTCTGACAGCCCTCCATTTGATGTATCCAGAGCTGGGCGAGAAAGATTCGG	2100
1019	ATCACCAACACCAACCAACCAACCAATCAACAGCAGCAGATCAGAGCCATCA	1078	2099	GATTTACTGATCTCCCAAGAGATCAGATTTACTTTGATTTGAGCTTTTGGAGC	2158
1021	ATCACCAACACCAACCAACCAACCAATCAACAGCAGCAGATCAGAGCCATCA	1080	2101	GATTTACTGATCTCCCAAGAGATCAGATTTACTTTGATTTGAGCTTTTGGAGC	2160
1079	TTCCTCAGCTCCAGCGGAGCAGGCTGTGCGGCGCTTCCCGCGAGCGGGCGTATGG	1138	2159	TGTTGTCTCAGATTTCCATCAGTCAACACCTGCTGAGATAAGTTTGTGTTCTGCA	2218
1081	TTCCTCAGCTCCAGCGGAGCAGGCTGTGCGGCGCTTCCCGCGAGCGGGCGTATGG	1140	2161	TGTTGTCTCAGATTTCCATCAGTCAACACCTGCTGAGATAAGTTTGTGTTCTGCA	2220
1139	AGTCCCAACCTCCACCCCAACCAACCGCGCTTCCCGCGAGCGGGCGGCTTATGG	1198	2219	ATGGAGTGTCTCTGATCGACTTTCAGTGCCTTTCGTGGATTTGGGAGTGGCTCGACTCTA	2278
1141	AGTCCCAACCTCCACCCCAACCAACCGCGCTTCCCGCGAGCGGGCGGCTTATGG	1200	2221	ATGGAGTGTCTCTGATCGACTTTCAGTGCCTTTCGTGGATTTGGGAGTGGCTCGACTCTA	2280
1199	ACGAGGACCTCGCTCGCGCGCGCTGCTACCGCGCGCTTCCCGCGAGCGGGCGG	1258	2279	TTAAAGACTTTTCTTAAATTTGAGAGCTTGCATATCCAGCTTAGCCTTCC	2338
1201	ACGAGGACCTCGCTCGCGCGCGCTGCTACCGCGCGCTTCCCGCGAGCGGGCGG	1260	2281	TTAAAGACTTTTCTTAAATTTGAGAGCTTGCATATCCAGCTTAGCCTTCC	2340
1259	TGAAGCGGCTCCCGCGCGCGCGCTTCCCGCTTCCCGCTTCCCGCTTCCCGCTTCC	1318	2339	TGTCAGCACTGAGATGATCAGAGAAAGACATGGGTAAAGAAACCAAGAGAGTCTGAG	2398
			2341	TGTACGACTGAGATGATCAGAGAAAGACATGGGTAAAGAAACCAAGAGAGTCTGAG	2400



QY 2399 AGCTATGCAACAGATCAACAGCAGTTTAAAGACACACAGAGTAAGGGACAGGCTCTGG 2458  
DB 2401 AGCTATGCAACAGATCAACAGCAGTTTAAAGACACACAGAGTAAGGGACAGGCTCTGG 2460  
QY 2459 AGCCACCGAGTCCAGGTCCTGGGTCCTGGTGAACCTGAGGAAGATCTCACCCCTGG 2518  
DB 2461 AACCCNACAGTCCAGGTCCTGGTGGCCCTGGTAGAATCTGAGGAAGATCTGACCCCTGG 2520  
QY 2519 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2578  
DB 2521 GCCTCAGCGCATCTTCTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTG 2580  
QY 2579 ACAAGCTCTTCTGGACACCCCTACCTTCTATACAGGACAGTGAGGAGTAGTGCTCT 2638  
DB 2581 ACAAGCTCTTCTGGACACCCCTACCTTCTATACAGGACAGTGAGGAGTAGTGCTCT 2640  
QY 2639 CTTCTCTAGCAGCTGCTTGTACGACGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2698  
DB 2641 CTTCTCTAGCAGCTGCTTGTACGACGCAAGAGGATAGTTTGGAAACCTATCATTTCC 2700  
QY 2699 TGTCTCTCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2758  
DB 2701 TGTCTCTCTTAAGAGGAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTCTG 2760  
QY 2759 GCTCTTTCTTCAACCTAAAGCCAGAAACCTTGACAGAGTATGTGTGGGTTGTGTT 2818  
DB 2761 GCTCTTTCTTCAACCTAAAGCCAGAAACCTTGACAGAGTATGTGTGGGTTGTGTT 2820  
QY 2819 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2878  
DB 2821 TTATATTAGGCAATTTGGGGATGGGGTGGAGGGGTATAGTTCATGAGGGTTTCTAA 2880  
QY 2879 GAAATGCTTAACAAAGCACTTTTGACAATGCTATCCACAGCAGGAAAAAGGATAATA 2938  
DB 2881 GAAATGCTTAACAAAGCACTTTTGACAATGCTATCCACAGCAGGAAAAAGGATAATA 2940  
QY 2939 TAATCTGTTTTAAACCTTTCTGGGGAATCCAAATATAGTGTGTTGTATTTAAACAA 2998  
DB 2941 TAATCTGTTTTAAACCTTTCTGGGGAATCCAAATATAGTGTGTTGTATTTAAACAA 3000  
QY 2999 GAACAGCCAGGTTGTCGCGAGGTTAGGATGTGTTTAAAGATTGGTCCCTTGAATAT 3058  
DB 3001 GAACAGCCAGGTTGTCGCGAGGTTAGGATGTGTTTAAAGATTGGTCCCTTGAATAT 3060  
QY 3059 ATGCTTCTGTATCAAGGTACGTATGTGTGTCGCAACAAAGGCGAGAACTTCTTTTAAAT 3118  
DB 3061 ATGCTTCTGTATCAAGGTACGTATGTGTGTCGCAACAAAGGCGAGAACTTCTTTTAAAT 3120  
QY 3119 TCCTTCTCTTTATTTTAAACAAATGGTGAAAGATGGAGGATTAACCAAAATCAGACAT 3178  
DB 3121 TCCTTCTCTTTATTTTAAACAAATGGTGAAAGATGGAGGATTAACCAAAATCAGACAT 3180  
QY 3179 GGCMAAACAAATATGCTGCTTGTCTTCCATTAACAGTGCATTTTAAAGTGTGCTCT 3238  
DB 3181 GGCMAAACAAATATGCTGCTTGTCTTCCATTAACAGTGCATTTTAAAGTGTGCTCT 3240  
QY 3239 TACTAAGTCTGTTTATTAACCTCTCTTTATCTATATGGAATAAAGAGGAGGAGTCA 3298  
DB 3241 TACTAAGTCTGTTTATTAACCTCTCTTTATCTATATGGAATAAAGAGGAGGAGTCA 3300  
QY 3299 TGTTAGCAATGACAGTTTAATATCCCTAGCAGAGCTGTGTTTACCTTCCCTGTGCGATC 3358  
DB 3301 TGTTAGCAATGACAGTTTAATATCCCTAGCAGAGCTGTGTTTACCTTCCCTGTGCGATC 3360  
QY 3359 CTTTCTGAGGTGCGCCATCCAGACTTTTAGGCCATTTCTGATGGAACCGATCCCTG 3418  
DB 3361 CTTTCTGAGGTATGCGCCATCCAGACTTTTAGGCCATTTCTGATGGAACCGATCCCTG 3420  
QY 3419 CCTGTAGTCTCAGCTATCTGAAAGTGGATCAGATTATAAAGTGAATTTACATGTAATCTG 3478  
DB 3421 CCTGTAGTCTCAGCTATCTGAAAGTGGATCAGATTATAAAGTGAATTTACATGTAATCTG 3480

QY 3479 TTTTGGTTGTGTTCTATCAACCCACAGAGTTCCCTAAACTTGTCTCAGTTATAGTAAC 3538  
DB 3481 TTTTGGTTGTGTTCTATCAACCCACAGAGTTCCCTAAACTTGTCTCAGTTATAGTAAC 3540  
QY 3539 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTTCAGTTCAGTATTTGATCCCTAGATAAG 3598  
DB 3541 TGACTGGTATATTCATTTCAGAAAGCCCATTAAGTTCAGTTCAGTATTTGATCCCTAGATAAG 3600  
QY 3599 AACATGCAAAATCAGCAGGAATCTGGTCATACAGGGTAAGCACAGGGACAATTAAGGATTTT 3658  
DB 3601 AACATGCAAAATCAGCAGGAATCTGGTCATACAGGGTAAGCACAGGGACAATTAAGGATTTT 3660  
QY 3659 TATAGTATATTAATTTTGTGTTA--TTGGTTAAGGAGA-CAATTTTGGAGGCAAGCA 3715  
DB 3661 TATAGTATATTAATTTTGTGTTAATTTTGGTTAAGGAGACCAATTTTGGAGGCAAGCA 3720  
QY 3716 AA---TCTTTTAAAAAATAGTATGAATGTGAATACTAGAAAAGATTAAAAAATAGTAT 3772  
DB 3721 AATCTTCTTTTAAAAAATAGTATGAATGTGAATACTAGAAAAGATTAAAAAATAGTAT 3780  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
DB 3781 GAGTGTGAGTACTAGGAAGGAT 3802

## RESULT 6

US-10-247-671-105  
; Sequence 105, Application US/10247671  
; Publication No. US20030194721A1  
; GENERAL INFORMATION:  
; APPLICANT: Mikita, Thomas  
; APPLICANT: Shiffman, Dov  
; APPLICANT: Porter, Gordon, J.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
; FILE REFERENCE: PA-0050 US  
; CURRENT APPLICATION NUMBER: US/10/247,671  
; CURRENT FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/323,784  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PERL Program  
; SEQ ID NO 105  
; LENGTH: 5828  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030194721A1 349343.3  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 1012-1074, 4044-4074, 5807  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-247-671-105

Query Match 97.0%; Score 3679.2; DB 16; Length 5828;

Best Local Similarity 98.1%; Pred. No. 0;

Matches 3728; Conservative 0; Mismatches 66; Indels 7; Gaps 4;

QY 1 ATAAATGAGTCCGAGAGCGAGCGACGCGAGCGAGCGGAGCGGAGTCTCCTGCGCT 60  
DB 1 ATAAATGAGTCCGAGAGCGAGCGAAACGCGAGCGCGGAGCGGAGTCTCCTGCGCT 60  
QY 61 CCGCGCCCCCACCCTCCAGCTCCTGCTCCTCCTCGCTCCCATACACAGAGCGCTCA 120  
DB 61 CCGCGCCCCCACCCTCCAGCTCCTGCTCCTCCTCGCTCCCATACACAGAGCGCTCA 120  
QY 121 CACCGCTCCCTCCTCCTGCGACACACAGACACAAGCGCGCACACAGGCTCG--CACAC 178  
DB 121 CACCGCTCCCTCCTCCTGCGACACACAGACACAAGCGCGCACACAGGCTCCGCA 180  
QY 179 ACTTGTCTCTCCGCGCGCTCACCCCTCTTGGCTGAGCCCTTGGCCGTGAGCGCG 238

181	Db	ACTTCGCTCTCCGCGCGCTCACACCCCTCTTGGCCCTGAGCCCTTGC CGGTGACGCGCG 240
239	Qy	QSCCGACGTGGACGCCCTCTCCGGGCTCACTTTGCAACGCTGACGTTGCCGCGCAGTGGC 298
241	Db	CGCCGACGTGGACGCCCTCTCCGGGCTCACTTTGCAACGCTGACGTTGCCGCGCAGTGGC 300
299	Qy	CGTGAAGTGGAAACAGCGGGGCGATCTCTCCGCTTGGTACAGGCCCAAGCCAGCAGCGCC 358
301	Db	CGTGAAGTGGAAACAGCGGGGCGATCTCTCCGCTTGGTACAGGCCCAAGCCAGCAGCGCC 360
359	Qy	CGCGAACTCTCGCTGTGTCTCTCCATGAGTCGGGATCGCAGATCCCCCACCAGCGCG 418
361	Db	CGCGAACTCTCGCTGTGTCTCTCCATGAGTCGGGATCGCAGATCCCCCACCAGCGCG 420
419	Qy	CTCACCGCTCCGGAGCGCTGGGCTTGTACACGCGAGCCCTTCGCGGACAGCAGCTGT 478
421	Db	CTCACCGCTCCGGAGCGCTGGGCTTGTACACGCGAGCCCTTCGCGGACAGCAGCTGT 480
479	Qy	GACTCCCCCAGTCGAGATTTCCGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 538
481	Db	GACTCCCCCAGTCGAGATTTCCGGACAGCTCTCTAGAACTCGCTCTAAAGACGGAAC 540
539	Qy	CGCCAAGCACTCAAAGCCCACTCGGGAAGAGGCGAGCCCGGCAAGCCCGGGCCCTGAGC 598
541	Db	CGCCAAGCACTCAAAGCCCACTCGGGAAGAGGCGAGCCCGGCAAGCCCGGGCCCTGAGC 600
599	Qy	CTGGACCTTAGGGTGGCGGCGACACTGCGCGGCTTCGCTCGCGGAGCTCCGCTC 658
601	Db	CTGGACCTTAGGGTGGCGGCGACACTGCGCGGCTTCGCTCGCGGAGCTCCGCTC 660
659	Qy	CTCTACACTCTCAGCCTCGCTGGAGAGACCCAGCGCCCACTTACAGCGCGCAAGAT 718
661	Db	CTCTACACTCTCAGCCTCGCTGGAGAGACCCAGCGCCCACTTACAGCGCGCAAGAT 720
719	Qy	ACCTTCAGANATGCCCTGGTCCAGCCCAATATAGCCCTTCCTCCAGTTCAGTT 778
721	Db	ACCTTCAGANATGCCCTGGTCCAGCCCAATATAGCCCTTCCTCCAGTTCAGTT 780
779	Qy	ATCGCGCGCAGACATACAGCTCGGAATACACACGGAGATCATGAACCCCGACTACCA 838
781	Db	ATCGCGCGCAGACATACAGCTCGGAATACACACGGAGATCATGAACCCCGACTACCA 840
839	Qy	AGTGAACATPAGACTTTGGAGACATCAGATCAGCGCTACAGCCACAGCTCCCTGCCA 898
841	Db	AGTGAACATPAGACTTTGGAGACATCAGATCAGCGCTACAGCCACAGCTCCCTGCCA 900
899	Qy	GCATCAGTACTTCTGGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCCTCGTGT 958
901	Db	GCATCAGTACTTCTGGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCCTCGTGT 960
959	Qy	ACCAATGACAGCGCCCTTGATCAAGTGAGAGGGGGCGGCCCGCAGTACCATCAC 1018
961	Db	ACCAATGACAGCGCCCTTGATCAAGTGAGAGGGGGCGGCCCGCAGTACCATCAC 1020
1019	Qy	ATCACCAACACCAACCAACCAACCATCACAGCAGCAGCATCAGCAGCATCA 1078
1021	Db	NN 1080
1079	Qy	TTCTCCAGCTCCAGCCCGGAGGACAGGTGTGCCCCAGCACTCATGTACTTCAAGC 1138
1081	Db	TTCTCCAGCTCCAGCCCGGAGGACAGGTGTGCCCCAGCACTCATGTACTTCAAGC 1140
1139	Qy	AGTCCCAACGCTCAACCCCAACACGCGGCTTCCTCCCGCAGGCGGGGGCGTATGG 1198
1141	Db	AGTCCCAACGCTCAACCCCAACACGCGGCTTCCTCCCGCAGGCGGGGGCGTATGG 1200
1199	Qy	ACGAGGCACTGCCCTCGCGGCCCGGCTGCATCGCACCCCGCGCTGTGGAACCGCGCGA 1258
1201	Db	ACGAGGCACTGCCCTCGCGGCCCGGCTGCATCGCACCCCGCGCTGTGGAACCGCGCGA 1260
1259	Qy	TGAAGCGGCTCCACAGTGGCGCGGCGCTTCGCGCTTCCTTCCACTTCAAGCCCTCGC 1318
1261	Db	TGAAGCGGCTCCACAGTGGCGCGGCGCTTCGCGCTTCCTTCCACTTCAAGCCCTCGC 1320

Qy	2397	AGAGCTATGCAACAAAGATCAACAGCATGTTTAAAGAACAACAGAGTAAAGGACAGGCTCT	2455
Db	2401	AGAGCTATGCAACAAAGATCAACAGCATGTTTAAAGAACAACAGAGTAAAGGACAGGCTCT	2460
Qy	2457	GGAGCCACCAAGGTCACAAAGTCCCTGGGTGCCCTGGTAGAAGCTCAGGGAAGATCTGCACCCT	2516
Db	2461	GGAGCCACCAAGGTCACAAAGTCCCTGGGTGCCCTGGTAGAAGCTCAGGGAAGATCTGCACCCT	2520
Qy	2517	GGGCTCCAGCGCATCTTCTACCTGAAAGCTGGAAGACTTGGTGTCTCCACCTTCCCATCAT	2576
Db	2521	GGGCTCCAGCGCATCTTCTACCTGAAAGCTGGAAGACTTGGTGTCTCCACCTTCCCATCAT	2580
Qy	2577	TGACAAAGCTCTTCTCGACACCTTACCTTCTTAATCAGAGAGCATGGAGCAGTGAGCTGC	2636
Db	2581	TGACAAAGCTCTTCTCGACACCTTACCTTCTTAATCAGAGAGCATGGAGCAGTGAGCTGC	2640
Qy	2637	CTCCTCTCCTAGCACCTGCTTGTCTACGCAAGAGGATAGGTTTGGAAACCTTATCATTTT	2696
Db	2641	CTCCTCTCCTAGCACCTGCTTGTCTACGCAAGAGGATAGGTTTGGAAACCTTATCATTTT	2700
Qy	2697	CCTGTCCTTCTTAAAGAGGAAAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTTTTTC	2756
Db	2701	CCTGTCCTTCTTAAAGAGGAAAAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTTTTTC	2760
Qy	2757	TGGCTCTTCTTCAACACCTAAAGCCAGAAAACTTCGACAGTATTGTGTGGGGTTGTG	2816
Db	2761	TGGCTCTTCTTCAACACCTAAAGCCAGAAAACTTCGACAGTATTGTGTGGGGTTGTG	2820
Qy	2817	TTTTATATTAGGCAATTTGGGGGATGGGGTGGGAGGGGGTTATAGTTTCATGAGGGTTTCT	2876
Db	2821	TTTTATATTAGGCAATTTGGGGGATGGGGTGGGAGGGGGTTATAGTTTCATGAGGGTTTCT	2880
Qy	2877	AAGAAAATTGCTTAAACAAAGCACTTTTGACAAATGCTATCCAGCAGGAAAAAAAAGGATAA	2936
Db	2881	AAGAAAATTGCTTAAACAAAGCACTTTTGACAAATGCTATCCAGCAGGAAAAAAAAGGATAA	2940
Qy	2937	TATAACTGTTTTAAAACTCTTTTCGGGAAATCCAATTTATAGTTGCTTTGTATTTAAAAAC	2996
Db	2941	TATAACTGTTTTAAAACTCTTTTCGGGAAATCCAATTTATAGTTGCTTTGTATTTAAAAAC	3000
Qy	2997	AAGAACAGCCAAAGGTTGTTGGCCAGGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAA	3056
Db	3001	AAGAACAGCCAAAGGTTGTTGGCCAGGGTAGGATGTGCTTAAAGATTGGTCCCTTGAAA	3060
Qy	3057	ATATGCTTCCTGTATCAAAAGTACGTATGTGGTGCAAAACAAAGGCAGAAAACTTCTTTTAA	3116
Db	3061	ATATGCTTCCTGTATCAAAAGTACGTATGTGGTGCAAAACAAAGGCAGAAAACTTCTTTTAA	3120
Qy	3117	TTTCCTTCTTCTTTTATTTTAAACAAATGGTGAAGATGGAGGATTACCTACAAATCAGAC	3176
Db	3121	TTTCCTTCTTCTTTTATTTTAAACAAATGGTGAAGATGGAGGATTACCTACAAATCAGAC	3180
Qy	3177	ATGGCAAAAACAATAATGGCTGTTTGCTTCCATAAACAAAGTGCAATTTTTTAAAGTGTCTGT	3236
Db	3181	ATGGCAAAAACAATAATGGCTGTTTGCTTCCATAAACAAAGTGCAATTTTTTAAAGTGTCTGT	3240
Qy	3237	CTTACTAAGTCTTGTTTTATTAACCTCTCTTTTATTTCTATATGGAAAAAAGAGGAGGCAGT	3296
Db	3241	CTTACTAAGTCTTGTTTTATTAACCTCTCTTTTATTTCTATATGGAAAAAAGAGGAGGCAGT	3300
Qy	3297	CATGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGCGA	3356
Db	3301	CATGTTAGCAAAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTCACCTTCCCTGTGCGA	3360
Qy	3357	TCCCTTCTGAGGATATGCCCCATCCAAGACTTTTAGGCCATTCTTGTATGGAAACAGATCCC	3416
Db	3361	TCCCTTCTGAGGATATGCCCCATCCAAGACTTTTAGGCCATTCTTGTATGGAAACAGATCCC	3420
Qy	3417	TGCCCTGACTGTCCAGCTATCTCCTGAAAAGTGATCAGATTATAAAGCTGATTTACATGTAAAC	3476
Db	3421	TGCCCTGACTGTCCAGCTATCTCCTGAAAAGTGATCAGATTATAAAGCTGATTTACATGTAAAC	3480
Qy	3477	TGTTTTGGTTGTGTTCTTATCAACCCCAACAGAGTTCCCTAAACTGCTTCAGTTTATAGTA	3536

Db	3481	TGTTTTGGTTGTCTTCTATCAACCCACCAGAGTTCCCTAAACTGCTTCAGTTAGTA	3540
Qy	3537	ACTGACTGGGTATATTCATTCCAGAAGCGCCATAGTTCAGTTGAGTATTTGATCCCTAGATA	3596
Db	3541	ACTGACTGGGTATATTCATTCCAGAAGCGCCATAGTTCAGTTGAGTATTTGATCCCTAGATA	3600
Qy	3597	AGAACATGCAAAATCAGCAGGAACCTGGTTCATACAGGGTAAAGCACCCAGGAGCAATAAGGATT	3656
Db	3601	AGAACATGCAAAATCAGCAGGAACCTGGTTCATACAGGGTAAAGCACCCAGGAGCAATAAGGATT	3660
Qy	3657	TTTATAGATATAATTTAAATTTTGGTTATTTGGTTAAAGGAGACAAATTTTGGAGACGACGAA	3716
Db	3661	TTTATAGATATAATTTAAATTTTGGTTATTTGGTTAAAGGAGACAAATTTTGGAGACGACGAA	3720
Qy	3717	A---TCTTTTAAAAAATAGTATGAAATGTGAATACTAGAAAAAGATTTTAAAAAATAGTATG	3773
Db	3721	ATCTTCTTTTAAAAAATAGTATGAAATGTGAATACTAGAAAAAGATTTTAAAAAATAGTATG	3780
Qy	3774	AGTGTGAGTACTAGGAAGGAT	3794
Db	3781	AGTGTGAGTACTAGGAAGGAT	3801
RESULT 7			
US-10-002-600-53			
; Sequence 53, Application US/10002600			
; Publication No. US20020137077A1			
; GENERAL INFORMATION:			
; APPLICANT: Hopkins, Christopher M.			
; APPLICANT: Peterson, David P.			
; APPLICANT: Cocks, Benjamin G.			
; APPLICANT: Hawkins, Phillip R.			
; TITLE OF INVENTION: GENES REGULATED IN ACTIVATED T CELLS			
; FILE REFERENCE: PA-0042 US			
; CURRENT APPLICATION NUMBER: US/10/002,600			
; CURRENT FILING DATE: 2001-10-25			
; PRIOR APPLICATION NUMBER: 60/243,521			
; PRIOR FILING DATE: 2000-10-25			
; NUMBER OF SEQ ID NOS: 116			
; SOFTWARE: PERL Program			
; SEQ ID NO 53			
; LENGTH: 5642			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Template ID: 349343.3			
; NAME/KEY: unsure			
; LOCATION: 1012-1074, 4045-4075			
; OTHER INFORMATION: a, t, c, g, or other			
US-10-002-600-53			
Query Match 96.7%; Score 3668.2; DB 13; Length 5642;			
Best Local Similarity 98.1%; Pred. No. 0;			
Matches 3728; Conservative 0; Mismatches 66; Indels 8; Gaps 5;			
Qy	1	ATAAATGACGTGCCGAGAGCGAGCGAAACGCGACCGGGAGAGCGGAGTCTCTTGCCT	60
Db	1	ATAAATGACGTGCCGAGAGCGAGCGAAACGCGACCGGGAGAGCGGAGTCTCTTGCCT	60
Qy	61	CCCGCCCCCACCCTCCAGTCTCTCTCTCGCTCCCATACACAGACGCGCTCA	120
Db	61	CCCGCCCCCACCCTCCAGTCTCTCTCTCGCTCCCATACACAGACGCGCTCA	120
Qy	121	CACCCGCTCCCTCACTCGCACACACAGACACAAGCGCGCACACAGGCTCCG--CACACAC	178
Db	121	CACCCGCTCCCTCACTCGCACACACAGACACAAGCGCGCACACAGGCTCCGACACACAC	180
Qy	179	ACTTGGTCTCCCGCGGTCTCACACCCCTTGTGCTTGTAGCCCTTGCCTGTGAGCGCGG	238
Db	181	ACTTGGTCTCTCCCGCGGTCTCACACCCCTTGTGCTTGTAGCCCTTGCCTGTGAGCGCGG	240



Db 2401 AAGAGCTATGCAACAGATCACAAGCAGTTTAAAGACCAACAGAGTAAGGGACAGGCTC 2460  
QY 2456 TGGAGCCACCGAGTCAAGGTCTCTGGTGCCCTGTGAGAACTGAGAAAGATCTGCACCC 2515  
Db 2461 TGGAGCCACCGAGTCAAGGTCTCTGGTGCCCTGTGAGAACTGAGAAAGATCTGCACCC 2520  
QY 2516 TGGGCTCCAGCGCATCTTCTACCTGAGCTGAAGACTTGGTGTCTCCACCTCCATCA 2575  
Db 2521 TGGGCTCCAGCGCATCTTCTACCTGAGCTGAAGACTTGGTGTCTCCACCTCCATCA 2580  
QY 2576 TTGACAAGCTCTTCTGGACACCTTACCTTCTTAATCAGGAGCAGTGGAGCAGTGAGCTG 2635  
Db 2581 TTGACAAGCTCTTCTGGACACCTTACCTTCTTAATCAGGAGCAGTGGAGCAGTGAGCTG 2640  
QY 2636 CTTCTCTCTAGCAGCTGCTGCTACGACGAAAGGATAGTGTGGAAACCTATCAT 2695  
Db 2641 CTTCTCTCTAGCAGCTGCTGCTACGACGAAAGGATAGTGTGGAAACCTATCAT 2700  
QY 2696 TCTGTCTCTCTTAAGGAGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTT 2755  
Db 2701 TCTGTCTCTCTTAAGGAGAAAGCAGCTCTGTAGAAAGCAAGACTTCTTTT 2760  
QY 2756 CTGGCTCTTCTTACAACCTTAAGCCAGAAACTTGCAGAGTATGTGTGGGCTGT 2815  
Db 2761 CTGGCTCTTCTTACAACCTTAAGCCAGAAACTTGCAGAGTATGTGTGGGCTGT 2820  
QY 2816 GTTTTATATTAGGCATTTGGGGATGGGGTGGAGGGGGTTATAGTTCATAGAGGTTTC 2875  
Db 2821 GTTTTATATTAGGCATTTGGGGATGGGGTGGAGGGGGTTATAGTTCATAGAGGTTTC 2880  
QY 2876 TAAGAAATTGCTTAACAAGCACTTTTGGACAATGCTATCCAGCAGCAAAAAGGATA 2935  
Db 2881 TAAGAAATTGCTTAACAAGCACTTTTGGACAATGCTATCCAGCAGCAAAAAGGATA 2940  
QY 2936 ATATACTGTTTAAACCTTTCTGGGAATCCAATTATAGTGTCTTGTATTTAAAAA 2995  
Db 2941 ATATACTGTTTAAACCTTTCTGGGNATCCAATTATAGTGTCTTGTATTTAAAAA 3000  
QY 2996 CAAGAACAGCAAGGTTGTTCGCCAGGTPAGATGTCTTAAAGATTGGTCCCTTGA 3055  
Db 3001 CAAGAACAGCAAGGTTGTTCGCCAGGTPAGATGTCTTAAAGATTGGTCCCTTGA 3060  
QY 3056 ATATGCTCTCTATCAAGGTATGCTATGTGTGCAACAGGCAAACTTCTTTTA 3115  
Db 3061 AATATGCTCTCTATCAAGGTATGCTATGTGTGCAACAGGCAAACTTCTTTTA 3120  
QY 3116 ATTTCTCTCTCTTATTTTAAACAATGGTGAAGATGGAGATTACCTTACAAATCAGA 3175  
Db 3121 ATTTCTCTCTCTTATTTTAAACAATGGTGAAGATGGAGATTACCTTACAAATCAGA 3180  
QY 3176 CATGGCAAAACAATAATGGCTGTTTGTCTTCCATPAAACAAGTGCATTTTTTAAAGTGCTG 3235  
Db 3181 CATGGCAAAACAATAATGGCTGTTTGTCTTCCATPAAACAAGTGCATTTTTTAAAGTGCTG 3240  
QY 3236 TCTTACTAGTCTGTGTTTATTACTCTCTTTTATTCTATATGCAATATAAGAGGCGAG 3295  
Db 3241 TCTTACTAGTCTGTGTTTATTACTCTCTTTTATTCTATATGCAATATAAGAGGCGAG 3300  
QY 3296 TCATGTTTAGCAATGACAGTTAATCTCCTAGCAGAGGCTGTGTTTCACTTCCCTGTGCG 3355  
Db 3301 TCATGTTTAGCAATGACAGTTAATCTCCTAGCAGAGGCTGTGTTTCACTTCCCTGTGCG 3360  
QY 3356 ATCCCTCTGAGTATGGCCATCCAGACTTTTAGGCCATTTCTGATGGAACAGATCC 3415  
Db 3361 ATCCCTCTGAGTATGGCCATCCAGACTTTTAGGCCATTTCTGATGGAACAGATCC 3420  
QY 3416 CTGCCCTGACTGTCAGCTATCTGAAAGTGGATCAGATTATAAATGGAATACATGTAA 3475  
Db 3421 CTGCCCTGACTGTCAGCTATCTGAAAGTGGATCAGATTATAAATGGAATACATGTAA 3480  
QY 3476 CTGTTTTGGTGTGTTCTATCAACCCACAGAGTTCCCTTAACTTGTCTTCACTATAGT 3535  
Db 3481 CTGTTTTGGTGTGTTCTATCAACCCACAGAGTTCCCTTAACTTGTCTTCACTATAGT 3540

QY 3536 AACTGACTGGTATATTCATTTCAGAGCGCCATTAAGTCAAGTTCAGTATTTTGTATCCCTAGAT 3595  
Db 3541 AACTGACTGGTATATTCATTTCAGAGCGCCATTAAGTCAAGTTCAGTATTTTGTATCCCTAGAT 3600  
QY 3596 AAGAATCATCAATCAGCAGGAACCTGGTTCATACAGGTTAAGCACCAGGCAATAAGGAT 3655  
Db 3601 AAGAATCATCAATCAGCAGGAACCTGGTTCATACAGGTTAAGCACCAGGCAATAAGGAT 3660  
QY 3656 TTTTATAGATATAAATTTAATTTTGTATTGTTTAAAGGAGACAAATTTTGGAGAGCAAGCA 3715  
Db 3661 TTTTATAGATATAAATTTAATTTTGTATTGTTTAAAGGAGACAAATTTTGGAGAGCAAGCA 3720  
QY 3716 AA---TCTTTTAAAAATAGTATGAATGTGAATCTAGAAAGATTTTAAAAATAGTAT 3772  
Db 3721 AATCTTCTTTTAAAAAATAGTATGAATGTGAATCTAGAAAGATTTTAAAGAAATAGTAT 3780  
QY 3773 GAGTGTGAGTACTAGGAAGGAT 3794  
Db 3781 GAGTGTGAGTACTAGGAAGGAT 3802

## RESULT 8

US-09-873-367C-229  
; Sequence 229, Application US/09873367C  
; Publication No. US20030165839A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; APPLICANT: Soppet, Daniel  
; APPLICANT: Endress, Gregory  
; APPLICANT: Augustus, Meena  
; APPLICANT: Ebner, Reinhard  
; APPLICANT: Carter, Kenneth  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
; FILE REFERENCE: 689290-64  
; CURRENT APPLICATION NUMBER: US/09/873,367C  
; CURRENT FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/236,842  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: U.S. 60/244,867  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: U.S. 60/245,084  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 1067  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 229  
; LENGTH: 4977  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-873-367C-229

Query Match 75.8%; Score 2875; DB 10; Length 4977;  
Best Local Similarity 98.1%; Pred. No. 0;  
Matches 3008; Conservative 0; Mismatches 40; Indels 19; Gaps 9;

QY 733 CCTGTGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTTCCAGTTATGCGG--CGCAGA 790  
Db 94 CTGTGCTCCAGGCCCAATATAGCCCTTCCCTCCAGGTTTCCAGTTATGCGGTCGTGAGA 153  
QY 791 CATACAGCTCGGAATACACCAACCGAGATCATGAACCCCGACTACACCAAGCTGACATGG 850  
Db 154 CATACAGCTCGGAATACACCAACCGAGATCATGAACCCCGACTACACCAAGCTGACATGG 213  
QY 851 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGTCCCTGCCAGCATCAGTACCT 910  
Db 214 ACCTTGGCAGCACTGAGATCAGGCTACAGCCACCAAGTCCCTGCCAGCATCAGTACCT 273  
QY 911 TCGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCTTCCGTGTACCAATGCAGC 970  
Db 274 TTGTGGAGGGCTACTCGAGCAACTACGAACTCAAGCTTCTTCCGTGTACCAATGCAGC 333

QY	971	GGCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACCATCACCAACACC	1030	QY	2051	TGACAGCCTCCATTGATGATATCCAGAGCTGGCGCAGAAAAGATTCCGGGATTTACTGATC	2110
Db	334	GGCCCTTGATCAAGTGGAGGAGGGCGGGCGCCAGCTACCATCACCATCACCAACACC	393	Db	1414	TGACAGCCTCCATTGATGATATCCAGAGCTGGCGCAGAAAAGATTCCGGGATTTACTGATC	1473
QY	1031	ACCAC	1090	QY	2111	TCGCCAAAGAAATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2170
Db	394	ACCAC	453	Db	1474	TCGCCAAAGAAATCAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1533
QY	1091	CAGACCCGAG	1150	QY	2171	GACTTCCATCAGGTCAACACATGCTGAAGATAAGTTGTTGTTCTGCAATGGAATGTTGTC	2230
Db	454	CAGACCCGAG	513	Db	1534	GACTTCCATCAGGTCAACACATGCTGAAGATAAGTTGTTGTTCTGCAATGGAATGTTGTC	1593
QY	1151	CCACCCCCAC	1210	QY	2231	TGATCCACTTCAGTGCCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTT	2290
Db	514	CCACCCCCAC	573	Db	1594	TGATCCACTTCAGTGCCTTCGTGGATTTGGGGAGTGGCTCGACTCTATTAAAGACTTTT	1653
QY	1211	CCTCGGGCCCGGCTGCATCGCACCCGCGCGCGCTGCTGGACCGCGGATGAAGCGGTCC	1270	QY	2291	CCTTAAATTTGCGAGGCTGAACTTGTATATCAAGCCTTAGCCTGCTGTAGACTGTA	2350
Db	574	CCTCGGGCCCGGCTGCATCGCACCCGCGCGCGCTGCTGGACCGCGGATGAAGCGGTCC	633	Db	1654	CCTTAAATTTGCGAGGCTGAACTTGTATATCAAGCCTTAGCCTGCTGTAGACTGTA	1713
QY	1271	CCACGGTGGCGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTGCGCGCGCATGCC	1330	QY	2351	GCATGATCACAAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACA	2410
Db	634	CCACGGTGGCGCGCGCGCTTCCCGCTCTTCCACTTCAAGCCCTGCGCGCGCATGCC	693	Db	1714	GCATGATCACAAGAAAGACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACA	1773
QY	1331	CCGCGCCAGCG	1390	QY	2411	AGATCACAAGCAGTTTAAAGACACACAGAGTGAAGGACAGGCTCTGGAGCCACCGAGT	2470
Db	694	CCGCGCCAGCG	753	Db	1774	AGATCACAAGCAGTTTAAAGACACACAGAGTGAAGGACAGGCTCTGGAGCCACCGAGT	1833
QY	1391	TCAGCCTGCGCGTGGAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1450	QY	2471	CCAGGTCTCTGGTGGCGCTGTAGAACTGAGGAAAGATCTGCACCTGGGCGCTCCAGCGCA	2530
Db	754	TCAGCCTGCGCGTGGAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	813	Db	1834	CCAGGTCTCTGGTGGCGCTGTAGAACTGAGGAAAGATCTGCACCTGGGCGCTCCAGCGCA	1893
QY	1451	ACCCTGACGGCTGCGCTGCGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1510	QY	2531	TCCTTCACTGAAGCTGGAAGACTTGGTGTCTCACCTTCATCATTTGACAGCTCTTCC	2590
Db	814	ACCCTGACGGCTGCGCTGCGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	873	Db	1894	TCCTTCACTGAAGCTGGAAGACTTGGTGTCTCACCTTCATCATTTGACAGCTCTTCC	1953
QY	1511	TCAGCCTGCGCGTGGAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1570	QY	2591	TGACACACCTTACCTTCTAATCAGAGCGTGAAGCAGTGAAGCTCTCTCTCTAGCA	2650
Db	874	TCAGCCTGCGCGTGGAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	933	Db	1954	TGACACACCTTACCTTCTAATCAGAGCGTGAAGCAGTGAAGCTCTCTCTCTAGCA	2013
QY	1571	CCAGCAGAGCTGCTGCTGCTGGCGAGGCGACGTGTGCGGTGCGGGGCAACCGCGCT	1630	QY	2651	CCTGCTCTCAGCAGCAAGGATAGTTTGGAAACCTATCATTTCTGCTCTCTCTTA	2710
Db	934	CCAGCAGAGCTGCTGCTGCTGGCGAGGCGACGTGTGCGGTGCGGGGCAACCGCGCT	993	Db	2014	CCTGCTCTCAGCAGCAAGGATAGTTTGGAAACCTATCATTTCTGCTCTCTCTTA	2073
QY	1631	GCCAGACTAGGGCTGCGAACCTGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGC	1690	QY	2711	AGAGAAAGCAGCTCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTGGCTCTTCCCT	2770
Db	994	GCCAGACTAGGGCTGCGAACCTGCGAGGGCTGCAAGGGCTTTTCAAGAGAACAGTGC	1053	Db	2074	AGAGAAAGCAGCTCTGTAGAAAGCAAGACTTTCTTTTCTGGCTCTGGCTCTTCCCT	2133
QY	1691	AGAAAAATGCAAAATATGTTGCTGGCAATAAAAACTGCCAGTAGACAAGAGACGTGC	1750	QY	2771	ACAACTTAAAGCCAGAAACTTGCAGAGTATTGTTGGGGTGTGTTTATATTTAGGC	2830
Db	1054	AGAAAAATGCAAAATATGTTGCTGGCAATAAAAACTGCCAGTAGACAAGAGACGTGC	1113	Db	2134	ACAACTTAAAGCCAGAAACTTGCAGAGTATTGTTGGGGTGTGTTTATATTTAGGC	2193
QY	1751	GAACCCGATGTCAGTATGTCAGTATTCAGAGTGTCTCAGTGTGGAAATGTTAAAGAG	1810	QY	2831	ATTGGGGATGGGGTGGGAGGGGTTATAGTTATAGGGTTCCTAAGAAATTCCTAAC	2890
Db	1114	GAACCCGATGTCAGTATGTCAGTATTCAGAGTGTCTCAGTGTGGAAATGTTAAAGAG	1173	Db	2194	ATTGGGGATGGGGTGGGAGGGGTTATAGTTATAGGGTTCCTAAGAAATTCCTAAC	2253
QY	1811	TTGTCCGTACAGATAGTCTGAAGGGAGGAGAGTGTCTGCTCTCCAAACCAAGAGCC	1870	QY	2891	AAAGCAGCTTTTGGCAAGTCTATCCAGCAGGAAAAAAGAGATAATACTGTTTTAA	2950
Db	1174	TTGTCCGTACAGATAGTCTGAAGGGAGGAGAGTGTCTGCTCTCCAAACCAAGAGCC	1233	Db	2254	AAAGCAGCTTTTGGCAAGTCTATCCAGCAGGAAAAAAGAGATAATACTGTTTTAA	2313
QY	1871	CATTACAAACAGAACTTCTCAGCCCTCTCCACTTCTCCTTCCAACTGCAATGATGATG	1930	QY	2951	AACTCTTTCTGGGGAAATCCAATTAATAGTTGCTTTGTAATTAATAAAGAGGCAAGG	3010
Db	1234	CATTACAAACAGAACTTCTCAGCCCTCTCCACTTCTCCTTCCAACTGCAATGATGATG	1293	Db	2314	AACTCTTTCTGGGGAAATCCAATTAATAGTTGCTTTGTAATTAATAAAGAGGCAAGG	2373
QY	1931	CCCTTGTCCAGCTTTAAGACTCAACACCCAGAGATCTTGATTTTCCAGATACCTGTC	1990	QY	3011	GTGTGTCGCGAGGTAGATGTCCTTAAAGATTTGGTCCCTTGAATAATGCTTCTGTA	3070
Db	1294	CTCTTGTCCAGCTTTAAGACTCAACACCCAGAGATCTTGATTTTCCAGATACCTGTC	1353	Db	2374	GTGTGTCGCGAGGTAGATGTCCTTAAAGATTTGGTCCCTTGAATAATGCTTCTGTA	2433
QY	1991	CCACTGACAGGCTGCTGCAGGCAAGATGCTGAGCATGTGCAACAATTTCTACAACTCC	2050	QY	3071	TCAGAGGTACGTATGTTGGTGCAAAACAGGCGAGAACTTCTTTTAAATTTCTTCTTCT	3130
Db	1354	CCACTGACAGGCTGCTGCAGGCAAGATGCTGAGCATGTGCAACAATTTCTACAACTCC	1413	Db	2434	TCAGAGGTACGTATGTTGGTGCAAAACAGGCGAGAACTTCTTTTAAATTTCTTCTTCT	2493
				QY	3131	TATTTTAAACAAATGTTGAAAGATGAGGATTAACCTACAAATCAGACATGGCAAAACAATA	3190







QY	1811	TTGTCGGTACAGATAGTCTGAAAGGGAGGAGAGGTCTGCTGCTTCCAAAACCAAGAGCC	1870
DB	1174	TTGTCGGTACAGATAGTCTGAAAGGGAGGAGAGGTCTGCTGCTTCCAAAACCAAGAGCC	1233
QY	1871	CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAAATCTGCATGATGAATG	1930
DB	1234	CATTACAAACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAAATCTGCATGATGAATG	1293
QY	1931	CCCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTATTTCCAGATACTGTC	1990
DB	1294	CTCTTGTCCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTATTTCCAGATACTGTC	1353
QY	1991	CCACTCACCAGGCTGCTGCAGGCACAGATGCTGAGCATGTGCACAACTTCTACAACTCC	2050
DB	1354	CCACTCACCAGGCTGCTGCAGGCACAGATGCTGAGCATGTGCACAACTTCTACAACTCC	1413
QY	2051	TGACAGCCTCCATTGATGTATCCAGAAAGCTGGGCAGAAAAGATTCGGGATTTACTGATC	2110
DB	1414	TGACAGCCTCCATTGATGTATCCAGAAAGCTGGGCAGAAAAGATTCGGGATTTACTGATC	1473
QY	2111	TCCCAAGAGAGATCAGACATTTACTTTTGAATCATGCCCTTTTGGAGCTGTTTGTCCCTCA	2170
DB	1474	TCCCAAGAGAGATCAGACATTTACTTTTGAATCATGCCCTTTTGGAGCTGTTTGTCCCTCA	1533
QY	2171	GACTTTCATCAGCTCAAAACACTGCTGAAAGATAAGTTTGTGTTTCTGCAATGGACTTGTC	2230
DB	1534	GACTTTCATCAGCTCAAAACACTGCTGAAAGATAAGTTTGTGTTTCTGCAATGGACTTGTC	1593
QY	2231	TGCATCGACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAGACTTTT	2290
DB	1594	TGCATCGACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAGACTTTT	1653
QY	2291	CCTTAAATTTGCAGAGCTCAACCTTGATATCAAAGCCTTAGCTCGCTCAGCACTGA	2350
DB	1654	CCTTAAATTTGCAGAGCTCAACCTTGATATCAAAGCCTTAGCTCGCTCAGCACTGA	1713
QY	2351	GCATGATCAGAGAAAGACATGGGTAAAGAAACCAAAGAGAGTFCGAAGAGCTATGCAACA	2410
DB	1714	GCATGATCAGAGAAAGACATGGGTAAAGAAACCAAAGAGAGTFCGAAGAGCTATGCAACA	1773
QY	2411	AGATCACAAGAGTTTAAAGACACACAGAGTAAAGGACAGGCTCTGGAGCCCAACGAGT	2470
DB	1774	AGATCACAAGAGTTTAAAGACACACAGAGTAAAGGACAGGCTCTGGAGCCCAACGAGT	1833
QY	2471	CCAAAGTCTCGGTCCTCGTGAACCTGAGGAAGATCTGCAACCTGGGGCTCCAGCGCA	2530
DB	1834	CCAAAGTCTCGGTCCTCGTGAACCTGAGGAAGATCTGCAACCTGGGGCTCCAGCGCA	1893
QY	2531	TCCTTCTACCTGAAGCTGGAAGACTTGTTGTCTCCACCTTCATCATTTGACAAAGCTTCC	2590
DB	1894	TCCTTCTACCTGAAGCTGGAAGACTTGTTGTCTCCACCTTCATCATTTGACAAAGCTTCC	1953
QY	2591	TGGACACCTACCTTCTTAATCAGGACAGTGGAGAGTGAAGTGCCTCCTCTCTAGCA	2650
DB	1954	TGGACACCTACCTTCTTAATCAGGACAGTGGAGAGTGAAGTGCCTCCTCTCTAGCA	2013
QY	2651	CCTGTTGCTACGACGACAAAGGATAGTTTGGAAAACTATCATTTCTGTCTCTCTTA	2710
DB	2014	CCTGTTGCTACGACGACAAAGGATAGTTTGGAAAACTATCATTTCTGTCTCTCTTA	2073
QY	2711	AGAGGAAAAGCAGCTCCCTGTAGAAGCAAGACATTTCTTTTTTTCTCTGGCTCTTTTCTCT	2770
DB	2074	AGAGGAAAAGCAGCTCCCTGTAGAAGCAAGACATTTCTTTTTTTCTCTGGCTCTTTTCTCT	2133
QY	2771	ACAACTTAAAGCCAGAAAACCTTGACAGATATTGTGTTGGGGTGTGTTTATATTTAGGC	2830
DB	2134	ACAACTTAAAGCCAGAAAACCTTGACAGATATTGTGTTGGGGTGTGTTTATATTTAGGC	2193
QY	2831	ATTGGGGGATGGGGTGGAGGGGGTTATAGTTTCATGAGGGTTTTCTAAGAAATTTCTTAAC	2890
DB	2194	ATTGGGGGATGGGGTGGAGGGGGTTATAGTTTCATGAGGGTTTTCTAAGAAATTTCTTAAC	2253

Qy	2891	AAAGCACTTTGGACAACTGCTATCCACGACGAAAAAAGGATAATAACTGTTTTAA	2950
Db	2254	AAAGCACTTTGGACAACTGCTATCCACGACGAAAAAAGGATAATAACTGTTTTAA	2313
Qy	2951	AACTCTTTCTGGGGAATCCAAATATAGTTGCTTGTATTTAAAAACAAGAACAGCAAGG	3010
Db	2314	AACTCTTTCTGGGGAATCCAAATATAGTTGCTTGTATTTAAAAACAAGAACAGCAAGG	2373
Qy	3011	GTTGTTCCGACGGTATAGGATGTCTTAAGAGATCGTCCCTGAAATATGCTTCTCTGTA	3070
Db	2374	GTTGTTCCGACGGTATAGGATGTCTTAAGAGATCGTCCCTGAAATATGCTTCTCTGTA	2433
Qy	3071	TCAAGAGTACGTATGTGTGCGAAAACAAGGCAGAAAATCTCTTTTAAATTCCTTCTCTT	3130
Db	2434	TCAAGAGTACGTATGTGTGCGAAAACAAGGCAGAAAATCTCTTTTAAATTCCTTCTCTT	2493
Qy	3131	TATTTTAACAAAATGGTGGAAGATGAGGATTAACCTACAATCAGACATGGCAAAACAATA	3190
Db	2494	TATTTTAACAAAATGGTGGAAGATGAGGATTAACCTACAATCAGACATGGCAAAACAAT	2551
Qy	3191	ATGCGCTTTGCTTCCATAAACAAAGTGCAATTTTTTAAAGTGCTGTCTTACTAAGTCTTG	3250
Db	2552	ATGCGCTTTGCTTCCATAAACAAAGTGCAATTTTTTAAAGTGCTGTCTTACTAAGTCTTG	2608
Qy	3251	TTTATTAACCTCTCCTTTATTTCTATATGGAATAATAAAGGAGCGCATGTTAGCAAAATG	3310
Db	2609	T---TATCTCTCTTATCTATATGCGGGAATAAATAAAGGAGCGCATGTTAGCAAAATG	2665
Qy	3311	ACAGTTAAATATCCCTAGCAGAGCGTGTGTTACCTTCCCTGTGCGATCCCTTCTGAGGTA	3370
Db	2666	ACAGTTAAATATCCCTAGCAGAGCGTGTGTTACCTTCCCTGTGCGATCCCTTCTGAGGTA	2725
Qy	3371	TGGCCCATCAAGAATTTTATAGGCCATTTGTATGGAAACAGATCCCTGCCCTGACTGCC	3430
Db	2726	TGGCCCATCAAGAATTTTATAGGCCATTTGTATGGAAACAGATCCCTGCCCTGACTGCC	2785
Qy	3431	AGCTATCCTGAAAGTGATCAGATATAAACTCGATTACATGTAACCTGTTTGGTGTGTG	3490
Db	2786	AGCTATCCTGAAAGTGATCAGATATAAACTCGATTACATGTAACCTGTTTGGTGTGTG	2845
Qy	3491	TCTATCAACCCCAACAGAGTTCCCTAACTTGCTTCAAGTTATAGTAACCTGCTGTATAT	3550
Db	2846	TCTATCAACCCCAACAGAGTTCCCTAACTTGCTTCAAGTTATAGTAACCTGCTGTATAT	2904
Qy	3551	TCATTCAGAGCGGCATTAAGTTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC	3610
Db	2905	TCATTCAGAGCGGCATTAAGTTCAGTTGAGTATTTGATCCCTAGATAAGAACATGCAATC	2959
Qy	3611	AGCAGGAATCGGTATACAGGGTAAGCAACAGGACAATAAGGATTTTTATAGATATAAT	3670
Db	2960	TCAGAGGACTGGTCATACAGGGTAAGCAACAGGACAATAAGGATTTTTATAGATATAAT	3019
Qy	3671	TTAAATTTTGTATTTGGTTAAGGAGACAATTTTGGAGAGCAACGAA---TCTTTTAAA	3727
Db	3020	TTAAATTTTGTATTTGGTTAAGGAGACAATTTTGGAGAGCAACGAAATCTCTTTTAAA	3079
Qy	3728	AAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAAAATAGTATGAGTGTGACTAG	3787
Db	3080	AAATAGTATGAATGTGAATCTAGAAAAAGATTTTAAAGAAATAGTATGAGTGTGACTAG	3139
Qy	3788	GAAGGAT 3794	
Db	3140	GAAGGAT 3146	

RESULT 10  
US-10-641-643-928  
; Sequence 928, Application US/10641643  
; Publication No. US20040077003A1  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; Susan G. Stuart  
; Jeffrey J. Seilhamer  
;

## TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641.643

FILING DATE: 14-Aug-2003

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: &lt;Unknown&gt;

FILING DATE: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 928:

SEQUENCE CHARACTERISTICS:

LENGTH: 2714 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g1311504

SEQUENCE DESCRIPTION: SEQ ID NO: 928 :

US-10-641-643-928

Query Match 70.9%; Score 2689.6; DB 17; Length 2714;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2692; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

552	AAAGCCCACTGCGGAGAGGAGCGCGGCAAGCCGCGGCGCTGAGCCTGGACCTTACG	611
439	AAAGCCCACTGCGGAGAGGAGCGCGGCAAGCCGCGGCGCTGAGCCTGGACCTTACG	498
612	GGTGCGGCGAGCACTGCGGCGCTTTCGCTGCGCGGAGCGTCCGCTCTCTACACTCTC	671
499	GGTGCGGCGAGCACTGCGGCGCTTTCGCTGCGCGGAGCGTCCGCTCTCTACACTCTC	558
672	AGCTTCGCTGGAGAGACCCCGAGCCCAATTCAGCGCGCAAGATACCTTCAAGATAT	731
559	AGCTTCGCTGGAGAGACCCCGAGCCCAATTCAGCGCGCAAGATACCTTCAAGATAT	618
732	GCCTCGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGAGAC	791
619	GCCTCGCTCCAGCCCAATATAGCCCTTCCCTCCAGGTTCCAGTTATGCGGCGAGAC	678
792	ATACAGCTCGGAATACACACCGAGATCATGAACCCCGAGCTACACCAAGCTGACCATGGA	851
679	ATACAGCTCGGAATACACACCGAGATCATGAACCCCGAGCTACACCAAGCTGACCATGGA	738
852	CCTTGGCAGCACTGAGATCAAGGCTACAGCCACCAAGTCCCTGCGCAGCATAGTACCTT	911
739	CCTTGGCAGCACTGAGATCAAGGCTACAGCCACCAAGTCCCTGCGCAGCATAGTACCTT	798
912	CGTGGAGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTGCGTGTACCAATGCGAGC	971
799	CGTGGAGGCTACTCGAGCAACTACGAACTCAAGCCCTTCTGCGTGTACCAATGCGAGC	858
972	GCCTTTGATCAAGTGGAGAGGCGGCGCGCCAGCTACCATCACCATCACCAACCA	1031
859	GCCTTTGATCAAGTGGAGAGGCGGCGCGCCAGCTACCATCACCATCACCAACCA	918
1032	CCACCAACCAACCAACCAATCACAGCAGCAGCATCAGAGCCATTCATTCCTCCAGCCTC	1091
919	CCACCAACCAACCAACCAATCACAGCAGCAGCATCAGAGCCATTCATTCCTCCAGCCTC	978
1092	CAGCCCGAGGAGAGGAGTGTCTGCGCAGCAGCTTCATGTACTTCAAGCAGTCCCGCAGCTC	1151
979	CAGCCCGAGGAGAGGAGTGTCTGCGCAGCAGCTTCATGTACTTCAAGCAGTCCCGCAGCTC	1038
1152	CACCCCGACAGCGCGGCTTCCCGCGCAGGCGGCGGCTTATGGAGCAGGAGCAGTCC	1211
1039	CACCCCGACAGCGCGGCTTCCCGCGCAGGCGGCGGCTTATGGAGCAGGAGCAGTCC	1098
1212	CTCGGCGCGCGCTGATCGCACCCCGCGCTGTGTGGACCCCGCCGATGAAGCGGTCC	1271
1099	CTCGGCGCGCGCTGATCGCACCCCGCGCTGTGTGGACCCCGCCGATGAAGCGGTCC	1158
1272	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1331
1159	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1218
1332	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1391
1219	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1278
1392	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1451
1279	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1338
1452	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1511
1339	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1398
1512	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1571
1399	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1458
1572	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1631
1459	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1518
1632	CAGCGTGGCGGCGCGGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCAGTCCCGC	1691



Db 540 CCGCAAGCCCGGCGCTGAGCCTGGACCCTCAACAGAGCGGGCCAGCACACGCGGCGC 599  
Qy 630 CGGCGCTTCGCTCGCGGAGCGTCCGCTCTCTCACTCTCAGCCTCCGCTGGAGAGAC 689  
Db 600 GGTGCTTGGCTATCCGACGTCGCGGCTCTCACTCTCAGCCTCCGCTGGAGAGAC 659  
Qy 690 CCCAGCCCAACATTACGCGGCAAGATACCTCCAGATATGCGCTGCTGCTCAAGCCCA 749  
Db 660 CCCAGCCCAACATTACGCGGCAAGATACCTCCAGATATGCGCTGCTGCTCAAGCCCA 719  
Qy 750 ATATAGCCCTTCGCTCCGCTTCAGGTTCCAGTTATGCGGCGCAGACATACAGCTCGGAATACAC 809  
Db 720 ATATAGCCCTTCGCTCCGCTTCAGGTTCCAGTTATGCGGCGCAGACATACAGCTTCGGAATACAC 779  
Qy 810 CAGGAGATCATGAACCCCGCATACACCAAGCTGACCATGAGCCTTGGAGCAGCTAGAT 869  
Db 780 CACAGAAATCATGAACCCCGCATATGCGGCGCAGATACCTCCAGATATGCGCTGCTGAGCAGCGGAT 839  
Qy 870 CAGGCTACAGCCACACACGTCCTGCGCCAGCATCAGTACCTTCGAGGAGGCTACATCGAG 929  
Db 840 CATGGCCAGCCGACAGCTCCCTGCGCCAGCTTCAGTACCTTCATGGAGGCTTACCCAG 899  
Qy 930 CAACTACGAATCAAGCCTTCCTGCTGTACCAAAATGCAG-----CGGCGCTT 977  
Db 900 CAGCTCGGAATCAAGCCTTCCTGCTGTACCAAAATGCCGCTTCTGGGCTTCGGCTTT 959  
Qy 978 GATCAAGTGGAGGAGGGGGCGCGAGCTACCATACCATACCAACACACACCA 1037  
Db 960 GATCAAGATGGAAGAGGGTCGCGAGCATGGCTACCAACACCAACCAACCATCACATCA 1019  
Qy 1038 CCACCAACCAATCACACAGCAGCATCAGCAGCCATCCATTCCTCAGCCTCCAGCCC 1097  
Db 1020 TATACCAACCAACCAAGCAG-----CAGCAGCGCTCATTCCTCTCCCTCTGGCCC 1073  
Qy 1098 GGAGGAGGAGGTGCTGCCAGCACCTCCATGTACTTCAAGCAGTCCCCACCGTCCACCCC 1157  
Db 1074 CGAGGAGGAGGTACTGCCAGCACCTCCATGTACTTCAAGCAGTCTCCGCGCTTAGCC 1133  
Qy 1158 CACACGCGGCTTCCCGCGCAGCGGGGGGGTTATGGAGAGGCACTGCGCTCGGC 1217  
Db 1134 GACCACTCAGGCTTCCCGCGCAGCGGGGGGGCTGTGGAGCAGAGTGGCTCTGC 1193  
Qy 1218 GCGCGCTCATCGCACCGGCGGCTGTGGACCGCGCATGAGCGGTCGCCAGGT 1277  
Db 1194 GCTGTGCTGATCGCTCCGCGACCGCTGTGGACCGCGCATGAGCGAGTGGCCCCAAT 1253  
Qy 1278 GCGCGCTGCGGCTTCCCGCTCTTCCACTTCAAGCCTTCGCGCGCATCCCGCGCGC 1337  
Db 1254 GCGCGCTGCGGCTTCCCGATCTTCTTCAAGCCTTACCGCGCACCTCCCGGCGC 1313  
Qy 1338 CAGCGCGCGCGGCGCACACCTCGGCTACGACCGCGCGCTGCGCGCTCAGCCT 1397  
Db 1314 CAGCCAGCGCGGCGGCACACCTGGGCTATGACCCGCGCGCAGCTGCGCTCAGTCT 1373  
Qy 1398 GCGCTGGAGCGGAGCGCGCGGCGGCGAGCGCGCGCGCTTGGAGGCAACCGTA 1457  
Db 1374 ACCCTTGGAGCGCGCGCGCGCGCGGCGAGCTGTGCGCTCGAGGCGCATCCGTA 1433  
Qy 1458 CGGCTGCGGCTGGCCCAAGAGCGCGCGCTGCGCTTCCCGCTCTCGGCTCAGCGC 1517  
Db 1434 CGGCTGCGGCTGGCCCAAGAGCGCGCGCTGCGCTTCCCGCTCGGCTCAGCGC 1493  
Qy 1518 CTCCTTACGCGCTCAGCCTGCTGGCGGAGAGTCCAGCCTGCGCTGCGCGCGCGAGCAG 1577  
Db 1494 GTCCCTTACGCGCTCAGCCTGCTGGGAGAGAGCGCGCAGCTACCATCGCACCCCAATAG 1553  
Qy 1578 GAGCTGCTGCTGGCGGAGCGCTGTGCGGTGGGGGCAACCGCGCTCCGAGCA 1637  
Db 1554 GAGCTCATCATCGCGGAGGCGCGCTGTGCTGTGCGGGGCAAACTGCTGCTGCGCAGCA 1613  
Qy 1638 CTACGCGGCTCGAAACCTGCGAGGCTGCAAGGCTTTTTCAGAGAACAGTGCAGAAAAA 1697  
Db 1614 CTACGAGTCCGACCTCGAGGGCTGCAAGGCTTCTTCAGAGAACCGTGCAGAAAAA 1673

Qy 1698 TGCATAATATGTTTGCCTGGCAAAATAAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG 1757  
Db 1674 CGCAAAATATGTTTGCCTGGCAAAATAAAAACTGCCAGTAGACAAGAGAGCTCGAAATCG 1733  
Qy 1758 ATGTCTAGTACTGTTCGATTTTCAAGAGTGTCTCAGTGTGGAATGGAATAAGAGTGTGTCG 1817  
Db 1734 ATGTCTAGTACTGTTCGAGTTTCAAGAGTGTCTCAGTGTGGAATGGAATAAGAGTGTGTCG 1793  
Qy 1818 TACAGATAGTCTGAAAGGAGGAGAGGTCTGTGCTTCCAAACCAAGAGCCCATTTACA 1877  
Db 1794 TACAGATAGTCTGAAAGGAGGAGAGGTCTGTGCTTCCAAACCAAGAGCCCATTTACA 1853  
Qy 1878 ACAGAACCTTCTCAGCCCTCTCCAGCTTCTCTCAATCTGCATGATGATGCCCTGT 1937  
Db 1854 ACAGGAGCCCTCGCAGCCCTCCCAACCTCTCTCGATCTGTATGATGAACCCCTGT 1913  
Qy 1938 CCGAGCTTTAAACAGACTCAACCCAGAGATCTTGATTTATCCAGATACTGTGCCACTGA 1997  
Db 1914 CCGAGCTTTAAACAGAGCAACGCCAGAGACCTTGATTTACTCCAGATACTGTGCCACGA 1973  
Qy 1998 CCGGCTGTGCGGCGACAGATGCTGAGCATGTGCAACAAATTTCTAACCTCTGACAGC 2057  
Db 1974 CCGAGCCACTGCGGCGACAGACGCTGAGCAGGTGCGAGAGTTCTACAACTTCTGACGCG 2033  
Qy 2058 CTCATTGATGATCCAGAGCTGGCAGAAAGATTCCGGGATTTACTGATCTCCCA 2117  
Db 2034 CTCCTATCGAGCTGTCAGAGCTGGCAGAAAGATCCCGGATTCAGTATCTCCCA 2093  
Qy 2118 AGAAGATCAGACATTAATTTGAATCAGCCTTTTGGAGCTGTTTGTCTCAGACTTTC 2177  
Db 2094 AGAAGATCAGAGCTTAATTTAGATCAGCCTTTTGGAGCTGTTGTTCTTAGACTTTC 2153  
Qy 2178 CATCAGGTCAAAACACTGTGAAAGATAAGTTGTGTTCTGCAATGGAAGTGTCTGATCG 2237  
Db 2154 TATCAGGTCAAAACACTGTGAAAGATAAGTTGTGTTCTGCAATGGAAGTGTCTGATCG 2213  
Qy 2238 ACTTCAGTCCCTCGTGGATTTGGGAGTGGCTCGACTCTATTAAGAGCTTTTCTTAAA 2297  
Db 2214 ACTTCAGTCCCTCGCGGATTTGGGAGTGGCTCGACTCCATTTAAAGAGCTTTTCTTAAA 2273  
Qy 2298 TTTGAGAGCCTGAACCTTTGATATCAAGCCTTTAGCTGCTGCTCAGCAGTGAAGTATGAT 2357  
Db 2274 TTTGAGAGCCTGAACCTTTGATATCAAGCCTTTAGCTGCTGCTCAGCAGTGAAGTATGAT 2333  
Qy 2358 CACAGAAAGATGAGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCAC 2417  
Db 2334 CACAGAGCAGATGGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACAGATCAC 2393  
Qy 2418 AAGCAGTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGCCCAACCGAGTCCAAGGT 2477  
Db 2394 AAGCAGTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGCCCTCAGAGCCCAAGGT 2453  
Qy 2478 CTTGGTGCCTTGTGTAACCTGAGGAAGATCTGCACCTGGGCTCTCAGCGCATCTTCTA 2537  
Db 2454 CTTTGGCGCAGCTGTTGGAATCTGAGGAAGATCTGCACCCAGGGCTCAGGGTATCTTCTA 2513  
Qy 2538 CTTGAAGCTGGAAGACTTGGTGTCTCCACCTTTCATCATTTGAACAAGCTTTCCTGAGCAC 2597  
Db 2514 CTTGAAGCTGGAAGACTTGGTGTCTCCACCTTTCATCATTTGAACAAGCTTTCCTGATAC 2573  
Qy 2598 CCTACCTTTCTAATCAGGAGCAG-TGGAGCAGTGAAGTGGCTCTCTCTAGCAGCTGCT 2656  
Db 2574 CTTGCTTTCTGAGCAGGGGAAGCCTGAGCAGAGACTTCTGCTCTGCTGGCAGCTGCT 2633  
Qy 2657 TGTACGCGCAGCAAGGATAGGTTTGGAAACCTTATCATTTTCTGCTCTTCTTAAAGAGGA 2716  
Db 2634 ATTAAGTGAACAAAGATGGGTTGAACACCT-GCCCTCTATCTTCTCTCAGGGGAA 2692  
Qy 2717 AAAGCAGCTCCTGTGAAAGCAAGACTTCTTTTTTTTCTGGCTCTTTTCTTACAAAC 2776  
Db 2693 AAAGCAGCTCCCATAGAAAGCAAGACTTT-TTTTTTCTGSCACCTTTCTTCTTACAAAC 2751



Db 720 ATATAGCCCTTCGCTCCGGGTCCACTTATGCCCAGCAGACTTATGGCTCGGAATACAC 779  
Qy 810 CACGGAGATCATGAACCCCGACTACACCAAGCTGACCATGGACCTTTGGCAGCAGCTAGAT 869  
Db 780 CACAGAAATCATGAACCCCGACTATGCAAGCTGACCATGGACCTCGGTAGCAGCGGAT 839  
Qy 870 CACGGCTACAGCCACACAGCTCCCTGCGCAGCATCAGTACTTGGTAGGGCTACTCGAG 929  
Db 840 CATGGCCACGGCCACAGCTCCCTGCGCAGCTTCAGTACTTCATGGAGGGCTACCCAG 899  
Qy 930 CAACTACGAACCTCAAGCCCTTCGCTGTACCAAAATGCAG-----CGGCCCTT 977  
Db 900 CAGCTCGGAACCTCAAGCCCTTCGCTGTACCAAAATGCCGCTTCGCGGCTCGGCCCTT 959  
Qy 978 GATCAAAAGTGGAGGCGGCGGCCAGCTACCATCAACCATCAACACACACACCA 1037  
Db 960 GATCAAGATGGAAGGGTGGGAGCATGGCTACCAACACACACACCATCAACATCA 1019  
Qy 1038 CCACACACCATCAACAGGAGCAGCATCAGAGCCATCGATTCCTCGAGCTCCAGCCC 1097  
Db 1020 TCATCACCAACCAACAGCAG-----CAGCAGCGCTCCATTCCTCCCTCTGGCCC 1073  
Qy 1098 GGAGGACGAGGTGCTGCCAGACCTCCATGTACTTCAAGCAGTCCCAACCGTCCACCCC 1157  
Db 1074 CGAGGACGAGGTACTGCCAGACCTCCATGTACTTCAAGCAGTCTCCGCGCTTAGGCC 1133  
Qy 1158 CACCACCGCGCTTCCCCCGCAGCGGGGGCGTTATGGGACGAGGCACTGCGCTCGGC 1217  
Db 1134 GACCACTCCAGGCTTCCCCCGCAGCGGGGGCGTGTGGAGCAGAGTGCCTCTGCG 1193  
Qy 1218 GCCCGCTGATCGCACCGCGCGCTGCTGAGCCCGCGATGAAGGGCGTCCACGGT 1277  
Db 1194 GCTTGGCTGATCGCTCCGGGACCGCTGCTGGACCGCGCAGATGAAGGCGAGTGCCTCAAT 1253  
Qy 1278 GGCGCGCGCGCTTCCCGCTTCCACTTCAAGCCCTCGCGCGCATCCCCCGCGCC 1337  
Db 1254 GGCGCTGCTGGCGCTTCCGATCTTCAAGCCCTCACCGCCACACCTCCCGCGCC 1313  
Qy 1338 CAGCCCGCGCGGCGCACCATCGGCTACGACCGAGCGCGCTGCGCGCTCAGCCT 1397  
Db 1314 CAGCCAGCGCGGCGGCCACCATCGGCTATGACCCCAACCGCGCAGCTGGCTCAGTCT 1373  
Qy 1398 GCGGCTGGAGCGCAGCGCGCGCGGAGCAGCGCGCGCTGAGAGCAGCCCGTA 1457  
Db 1374 ACCCTTGGAGCGCGCGCGCGCGGCGAGCAGCTGCTGCGCTCGAGGGCGCATCCGTA 1433  
Qy 1458 CGGCTGCGCTGGCCAGAGGCGCGCGCGCTGGGCTTCCCGCTCAGCGCTCAGCGC 1517  
Db 1434 CGGCTCCGCTGGCCAGAGGAGCGCGCAGCTTGAACCTTCCCTCCGCTGGGCTCAGCG 1493  
Qy 1518 CTCCTTACCGCTCCAGCTGCTGGGCGAGAGTCCAGCTGCGCTGCGCGGCCAGAGC 1577  
Db 1494 GTCCCTTACCGCTCCAGCTGCTGGGAGAGAGCCCGCTACCATCGCCACCAATAG 1553  
Qy 1578 GAGCTCGTCTGGGCGAGGCGACGTGTCGCTGGGGGAGCAACCGCGCTTCGAGCA 1637  
Db 1554 GAGCTCATATCCGCGGAGGCGACGTGTGTGTGTGGGGGACAAATGCTGCTGCCAGCA 1613  
Qy 1638 CTACGGCGTGGAACTCGGAGGCTGCAAGGCTTTTCAAGAGAACAGTGCAGAAAA 1697  
Db 1614 CTACGGAGTCCGACCTGGAGGGCTGCAAGGCTTCTTCAAGAGAACGGTGCAGAAAA 1673  
Qy 1698 TGCAAAATATGTTGCTGGCAATAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG 1757  
Db 1674 CGCAAAATATGTTGCTGGCAATAAAACTGCCCGGTAGACAAGAGAGCTCGAAATCG 1733  
Qy 1758 ATGTAGTACTGTGCAATTTCAAGAGTGTCTCAGTGTGGAATGGTAAAAAGAGTGTCCG 1817  
Db 1734 ATGTAGTACTGTGCAATTTCAAGAGTGTCTCAGTGTGGAATGGTAAAAAGAGTGTCCG 1793  
Qy 1818 TACAGATAGCTGAAGGGAGGAGGTGCTGCTCCCAACCAAGAGCCATTACA 1877

Db 1794 TACAGATAGCTGAAGGGAGGAGAGGTGCTGCTCCCTCCAAACCAAGAGCCACTACA 1853  
Qy 1878 ACAGGAACCTTCTCAGCCCTCTCCACCTTCTCCTCAATCTCGATGATGAATGCCCTTGT 1937  
Db 1854 ACAGGAGCCCTCGCAGCCCTCCACCATCTCCTCGATCTGTATGATGAAGCCCTTGT 1913  
Qy 1938 CCGAGCTTTAAACAGACTCAACCCAGCAGATCTTGATATTTCCAGATCTGTGCCACTGA 1997  
Db 1914 CCGAGCTTTAAACAGAGCGAACGCCAGAGACCTTGATTTACTCCAGATCTGTCCCCACGA 1973  
Qy 1998 CCGAGCTGTGAGGCGACAGATGCTGAGCATGTGCAAAATTTCTACAACCTCTCGACGC 2057  
Db 1974 CCGAGCCACTGCGGGCAGACAGCGCTGAGCAGCGTGCAGCAGTCTTACAACCTCTGAGCGC 2033  
Qy 2058 CTCCATGTATGTATCCAGAGCTGGGCGAGAAAGATTCGGGATTTACTGTATCTCCCAA 2117  
Db 2034 CTCCATCGAGCTGTCCAGAAAGCTGGGCGAGAAAGATTCGGGATTTCACTGATCTCCCAA 2093  
Qy 2118 AGAAGATCAGACATTAATTTGATCAGCCCTTTTGGAGCTGTTGTCTCAGACTTTC 2177  
Db 2094 AGAAGATCAGAGCTTACTTTATAGAATCAGCCCTTTTGGAGCTGTTGCTTTAGACTTTC 2153  
Qy 2178 CATCAGGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTCGATCG 2237  
Db 2154 TATCAGTCAAAACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTCGACCG 2213  
Qy 2238 ACTTCAGTCCCTTGTGGATTTGGGAGTGGCTCGACTCTATTAAGAAGCTTTTCTTTAAA 2297  
Db 2214 ACTTCAGTCCCTTGGGATTTGGGAGTGGCTCGACTCCATTAAGAAGCTTTTCTTTAAA 2273  
Qy 2298 TTTTCAGAGCTTGAACCTTGATATCCAGCCCTTAGCTGCTGCTCAGCAGTGAAGTATGAT 2357  
Db 2274 TTTTCAGAGCTTGAACCTTGATATCCAGCCCTTAGCTGCTGCTCAGCAGTGAAGTATGAT 2333  
Qy 2358 CACAGAAAGACATGGGTTAAAAAGAACCAAGAGAGTCGAAGAGCTATGCAAAAGATCAC 2417  
Db 2334 CACAGAGCAGATGGGTTAAAGAACCAAGAGAGTCGAAGAGCTATGCAAAAGATCAC 2393  
Qy 2418 AAGAGTTTTAAAGACCAACAGAGTAAGGACAGGCTCTGGAGCCCAACAGGTCCAGGT 2477  
Db 2394 AAGAGCTTTAAAGGACCAACAGAGGAGGACAGGCTCTGGAGCCCTCAGAGGCCAAGGT 2453  
Qy 2478 CTTGGGTGCTGTTAGAACTCAGGAGAGTCTGCACCTTGGGCTCCAGGCACTTCTA 2537  
Db 2454 CTTTCGCGCCTGTTGGAACTGAGGAAGATCTGCACCCAGGGCTCCAGGCTATCTCTA 2513  
Qy 2538 CTTGAAGCTTGAAGACCTTGTCTCCACCTTCCATCATTTGAACAAGCTCTTCTCGACAC 2597  
Db 2514 CTTGAAGCTTGAAGACCTTGTGTCCCCACCTTCTGTCTCAGACAGCTCTTCTCTGATAC 2573  
Qy 2598 CTTAOCCTTTCTAATCAGGAGCAG-TGGAGCAGTGAAGTGCCTCTCTCTAGACCTGCT 2656  
Db 2574 CTTGCTTCTGAGCAGGGGAGCTTGAACACCTTGCCTCTGCTGCGCAGCTGCTC 2633  
Qy 2657 TGTACCGAGCAAGGGATAGTTTGGAAACCTATCATTTCTGCTCTTCTTAAGAGGA 2716  
Db 2634 ATTAAGTGAAGCAAAAGGATGGGTTGAACACCTTGCCTCTCTTCTCAGGGGAA 2692  
Qy 2717 AAAGCAGCTCTGTGAAGAGCAAGACTTTCTTTTTTTTCTGGCTCTTTTCTTCAACACC 2776  
Db 2693 AAAGCAGCTCTTCAAGAGCAAGAGCAAGCTTT-TTTTTTCTGGCAGCTTCTTCAACACC 2751  
Qy 2777 TAAAGCAGAAACTTGCAGAGTATTTGTTGGGTTGTTTTATATTATTTAGGCAATGGG 2836  
Db 2752 TAAAGCAGAAACTTGCAGAGTATTTGTTGGGTTGTTTTATATTATTTAGGCTTTGGT 2811  
Qy 2837 GATGGGTTGGAGGGGGT--TATAGTTTCAAGGGTTTTTCTTAAGAAATGCTTAACAG 2894  
Db 2812 GGGTGGGCTGGAGGGGGTAAATAGTTTCATGAGGCTTTTCTTAAGAAATTTGCTGAGGAAG 2871  
Qy 2895 CACTTTTGGACAAATGCTATCCAGCAG-----GAAAAAAGGATAAATACTGTTTT 2948  
Db 2872 CACTTTTGGATGATGCTATCCAGCAGTGGGGTGGGGAGGAGGATATATACTGTTTT 2931











Qy	978	GATCAAAAGTGGAGGAGGGGGCGCGCCAGCTTACCATCACCAACCAACCAACCA	1037
Db	960	GATCAAGATGGAAGAGGGTGGCGAGCATGGCTACCAACACACACACCATCACCATCA	1019
Qy	1038	CCACCACCACTCACGAGCAGCAGCATCAGCAGCCCATCATTTCTTCCACGCTTCCAGCCC	1097
Db	1020	TCATCACACCAACACAGCAG-----CAGCAGCGTCCATTTCTCTCTCTGTGGCCC	1073
Qy	1098	GGAGGACGAGGTGCTGCCACGACCTCCATATGTACTTCAAGCAGTCCCACACCGTCCACCCC	1157
Db	1074	CGAGGACGAGGTACTGCGCCACGACCTCCATATGTACTTCAAGCAGTCTCCGCGGTCTAGGCC	1133
Qy	1158	CACCAACGCGGCTTTCCCGCGCAGGCGGGGGGGTTATGGGACGAGGCACTGCGCTCGGC	1217
Db	1134	GACCACTCCAGGCTTCCCGCGCAGGCGGGGGCGGTGTGGGACGACGAGTGGCTCTGCG	1193
Qy	1218	GCCCGGTGCATCGCACACCGCGCGCTGTGTGGACCCCGCCGATGAAGCGCGTCCCCACGGT	1277
Db	1194	GCCTGGGTGCATCGCTCGGGACCGCTGTCTGGACCCCGCAGATGAGGCACTGCGCCCAT	1253
Qy	1278	GGCGCGCGCGGCTTCCCGCTCTTCCACTTCAAGCCCTCGCCCGCGCATCCCCCGCGCC	1337
Db	1254	GGCGCTGCTGCGCGCTTCCCGATCTTCTTCAAGCCCTCACCGCACACCTCCCGCGCC	1313
Qy	1338	CAGCCCGCGCGGCGCACCACTCGGCTACGACCCGACGGCGCTGCGCGGCTCAGCCT	1397
Db	1314	CAGCCACGCGCGGGCGCACCACTTGGCTATGACCCACCGCGCGCAGCTGCGCTCAGTCT	1373
Qy	1398	GCCGCTGGGAGCCGACGCGCGCGGCGCAGCAGGCGCGCGCTTGGAGGCGCACCCGTA	1457
Db	1374	ACCCTTGGGAGCGCGCGCGCGGAGCGCAAGCTGTCTGCGCTCGAGGGCCATCGGTA	1433
Qy	1458	CGGCTGCCCTGGCCAAAGAGGGCGGCCCGCTGGCCCTTCCCGCTCTCGGCTCAAGCC	1517
Db	1434	CGGCTCCCGCTGGCCAAAGAGCGGCCACGTTGACCTTCCCTCGGTGGGCTCAGCAGC	1493
Qy	1518	CTCCCTTACCGCTCAGCCTGTGTGGCGGAGAGTCCAGCCTGCGGTGCGCGGCCAGCAG	1577
Db	1494	GTCCCTTACCGCGCTCAGCCTGTGTGGGAGAGAGCCCGAGCTTACCTCGGCCACCCAAATAG	1553
Qy	1578	GAGCTCGTCTGTGGCGAGGCGACGTGTGCGCTGTGGGGGACAAACGCCCTGCCAGCA	1637
Db	1554	GAGCTCATCATCGGCGAGGGCAGTGTGTGTGTGGGGGACAAATGCTGCTGCCAGCA	1613
Qy	1638	CTACGGCGTGGGAACCTTGGAGGGCTGCAAGGGCTTTTTTCAAGAGAACAGTCGAGAAAA	1697
Db	1614	CTACGGAGTCCGCACTTCGAGGGCTTCAAGGGGCTTCTTCAAGAGAACGGTGCGAGAAAA	1673
Qy	1698	TGCAAAATATGTTTTCCTGTGCAAAATAAAACTGCCAGTAGACAAGAGAGCTCGAAACCG	1757
Db	1674	CGCAAAATATGTTTGTGTGGCAATAAAACTTGCCCGGTAGACAAGAGAGCTCGAAATCG	1733
Qy	1758	ATGTCAGTACTGTTCGATTTTCAGAGTGTCTCAGTGTGTGAATGGTAAAGAAAGTTGTCCG	1817
Db	1734	ATGTCAGTACTGCAGGTTTCAGAAAGTGTCTCAGTGTGCGGATGGTGAAGGAAGTTGTGCG	1793
Qy	1818	TACAGATAGTCTGAAAGGAGGAGAGGTGCTGTGCTTCCAAACCAAGAGGCCCATTTACA	1877
Db	1794	TACAGATAGTCTGAAAGGAGGAGAGGTGCTGTGCTTCCAAACCAAGAGGCCCATTTACA	1853
Qy	1878	ACAGGAACCTTCTCAGGCCCTCTCCACCTTCTCTCCCAATCTGCATGATGAATGCCCTTGT	1937
Db	1854	ACAGGAGCCCTCGCAGGCCCTCCCCACCATCTCTCTCCGATCTGTATGATGAACGCCCTTGT	1913
Qy	1938	CCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTATTTCCAGATACTGTGCCACTGA	1997
Db	1914	CCGAGCTTTAACAGACGCAACGCCCCAGAGACCTTGATTTACTCCAGATACTGTCCACCGA	1973
Qy	1998	CCAGGCTGTGCGAGGACAGATGCTGAGCATGTGCAACAAATTTCTAACTCTCTGACG	2057
Db	1974	CCAGGCCACTGCGGGCACAGACGTGAGCACGTGACAGCTTCTACAACTTCTGACGGC	2033

QY	2058	CTCCATTGATGTATTCAGAAAGCTGGGCAGAAAAGATTCGGGATTTTACTGATCTCCCCAA	2111
Db	2034	CTCCATCGAGCTGTCCAGAAAGCTGGGCAGAAAAGATCCCGGATTCACCTGATCTCCCAA	2093
QY	2118	AGAAAGATCAGACATTACTTATCGAATCAGACCTTTTGGAGCTGTGTTGCTCTCAGACTTTC	2177
Db	2094	AGAAGATCAGAGCTTACTTATAGATCAGACCTTTTGGAGCTGTGTTGCTTATAGACTTTC	2153
QY	2178	CATCAGGTCAAAACACTGCTCAAGATAAAGTTTGTCTTGCAAATGGACTTGTCTCTGCATCG	2237
Db	2154	TATCAGGTCAAAACACTGCTCAAGATAAAGTTTGTGTTCTGCAATGGACTTGTCTCTGCACCG	2213
QY	2238	ACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTTAAA	2297
Db	2214	ACTTCAGTGCCTTCGTGGATTTTGGGAGTGGCTCGACTCTCAATTAAGAGACTTTTCTTTAAA	2273
QY	2298	TTTGCAGAGCTGAACCTTCATATCAAGCTTTAGCTGCTGCTGACGACTGAGCATGAT	2357
Db	2274	TTTGCAGAGCTGAACTTCATATCCAAAGCTTTAGCTGCTGCTGACGACTGAGCATGAT	2333
QY	2358	CACAGAAAGACATGGGTTTAAAGAACCAAAGAGAGTCGAAGAGCTATGCAACAAAGATCAC	2417
Db	2334	CACAGAGCAGATGGGTTTAAAGAACCAAAGAGAGTGGAGGAGCTATGCAACAAAGATCAC	2393
QY	2418	AAGCAGTTTAAAGACACACAGAGATTAAGGACAGGCTCTGGAGGCCACCGATCCAAAGT	2477
Db	2394	AAGCAGCTTAAAGACACACACAGAGAGGGAAGGCTCTGGAGCCCTCAGAGGCCCAAGGT	2453
QY	2478	CCTGGGTGCCCTGCTAGAACTGAGGAGAGATCTCACCCCTGGGCTCCAGCGCATCTTCTA	2537
Db	2454	CTTTCGGGCACTGGTGAACCTGAGGAAGATCTCGACCCAGGGCTCCAGGGTATCTTCTA	2513
QY	2538	CCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCATCATTTGACAAGCTCTTCTCGACAC	2597
Db	2514	CCTGAAGCTGGAGACTTGGTGTCCCCACCTCTGTGATCGACAAGCTCTTCTTGATAC	2573
QY	2598	CTTACTTTTCTAATCAGAGCAG - TGAGCAGTGAGCTGCTCCTCTCTAGCACCTGCT	2656
Db	2574	CCTGCCTTCTGAGCAGGGGAAGCTGAGCAGAGAGCTACTTGTCTCTGCTGGCACCTGGTC	2633
QY	2657	TGCTACGCAGCAAGGATAGGTTTGGAAACCTATCATTTCTCTGCTCTCTTAAGAGGA	2716
Db	2634	ATTAAAGTGCAAAAGGATGGGTTGAAACCT - GCCCTCTATATCTTCTCCAGGGGAA	2692
QY	2717	AAAGCAGCTCCTGTAGAAAGCAAGACTTCTTTTTTTTCTGCTCTTTTTCTTACAACC	2776
Db	2693	AAAGCAGCTCCATAGAAAGCAAGACTTT - TTTTTTCTGGCACCTTCTTACAACC	2751
QY	2777	TAAAGCCAGAAAACCTTCAGAGTATTGTGTGGGGTGTGTTTATATTAGGATTTGGG	2836
Db	2752	TAAAGCCAGAAAACCTTCAGAGTATTGTGTGGGGTGTGTTTATATTAGGCTTTGGT	2811
QY	2837	GGATGGGTCGGAGGGGT - TATAGTTTCATGAGGTTTCTAAGAAATGCTTAAACAAG	2894
Db	2812	GGGTGGCTCGGAGGGGGTAAATAGTTCATGAGGCTTTTCTAAGAAATGCTGACGAAG	2871
QY	2895	CACTTTTTGGACAATGCTATCCAGCAG - -----GAAAAAAGGATAATAACTGTTTT	2948
Db	2872	CACTTTTGGATGATGCTATCCAGCAGTGGGTTGGGAGAAAGGATAATAACTGTTTT	2931
QY	2949	AAAA - CTCCTTTCTGGGGAATCCAAATPATAGTTGCTTTGTATTATAAAAAACAAGACGCC	3006
Db	2932	AAAAACTCTTTTCGGGGGAATATGACTATGTTGCTTTGTATTATAAAAAAAGAACAGGCC	2991
QY	3007	AAGGGTTG - TTCGCCAGGGTAGGATGTCTTAAAGATTGGTCCCTTGAANAATATGCTTC	3065
Db	2992	AAGGGCTGTTTTTACCAGGGTAGGGCTGTGTCTTAAAGACTGATCCCTTTAGTATGTACTTC	3051
QY	3066	CTGTATCAAAAGGTACGTATGTTGGTGTGCAAAACAAGGCAGAAA - -----CTTCTCTT	3113
Db	3052	CCGATC - GAGGCACATAGTGGTGTGCAAAATGAGCGGGGGAATTCCTCATTTCTTCATTT	3110
QY	3114	TAATTTCTCTCTCTTTATTTTAAACAAATGGTGAAGATGGGAGGATTAACCTACAAATCA	3173





Search completed: March 22, 2005, 01:21:50  
Job time : 2024 secs

---

GenCore version 5.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 21, 2005, 08:40:51 ; Search time 11544 Seconds  
(without alignments)  
12510.034 Million cell updates/sec

Title: US-10-608-863-1  
Perfect score: 3794  
Sequence: 1 ataaatgcgcgcgcagaga.....gtgtgagtactaggagagat 3794

Scoring table: IDENTIFY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gssi: \*  
9: gb\_gse2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	879.8	23.2	996	1	AJ814207
2	796	21.0	1087	4	BM557376
3	769.4	20.3	1032	4	BM549822
4	752.4	19.8	819	5	BO431513
5	696.4	18.4	852	7	CK778210
6	627.8	16.5	671	6	CB445292
7	617	16.3	757	6	CD521819
8	612.2	16.1	679	6	CB465379
9	597.2	15.7	635	6	CB447412
10	593	15.4	940	5	BU436288
11	563	14.8	635	6	CD638747
12	561	14.8	794	6	CD522275
13	556.2	14.7	689	6	CB447773
14	555.4	14.6	802	7	CK635592
15	552.6	14.6	644	6	CB445564
16	537	14.2	735	6	CD523420
17	524.4	13.8	713	7	CF536097
18	520.4	13.7	586	5	BX925243
19	518.2	13.7	708	6	CD349619
20	495.4	13.6	533	2	BF086936
21	495.4	13.1	1797	9	AY417753
22	490.8	12.9	1797	9	AY417755
23	488.8	12.9	588	2	AW491078
24	480.8	12.7	540	5	BX920563

25	474.8	12.5	574	1	AI145834
26	474.4	12.5	2759	3	AK087495
27	461.6	12.2	493	2	BF604409
28	442.2	11.7	478	7	CR548239
29	429.4	11.3	537	7	N40404
30	428.8	11.3	441	1	AI911092
31	426.4	11.2	433	4	BG235965
32	425.6	11.2	1699	9	AY417754
33	422.8	11.1	525	2	BB758614
34	414.4	10.9	564	2	BF019466
35	411	10.8	500	2	AW490389
36	390	10.3	390	1	AV645408
37	390	10.3	390	1	AV645902
38	385.8	10.2	452	7	CR555348
39	376.4	9.9	483	6	CB441272
40	373.2	9.8	487	4	BI329884
41	372.6	9.8	451	1	AI327473
42	372.2	9.8	577	2	BE648145
43	365.8	9.6	407	5	BQ326341
44	363.8	9.6	2435	3	CR590328
45	351.2	9.3	428	1	AI323943

## ALIGNMENTS

RESULT 1  
AJ814207  
LOCUS AJ814207 KN206 Bos sp. cDNA clone C0005204p19, mRNA linear EST 13-SEP-2004  
DEFINITION AJ814207  
ACCESSION AJ814207  
VERSION AJ814207.1 GI:51881683  
KEYWORDS EST.  
SOURCE Bos sp.  
ORGANISM Bos sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
McGuire, K. and Glass, E.J.  
Unpublished ESTs, McGuire and Glass  
Unpublished (2004)  
Contact: McGuire K  
Genomics and Genetics  
Roslin Institute  
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM  
Single pass sequencing. Bases called and trimmed with phred v0.020425 c. Vector identified by cross match with the -minscore 20 and -minmatch 12 options. Vector: BluescriptII(SK+) R. Site 1: EcorV(lost) R. Site 2: NotI Seq Primer: T7 Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata.

FEATURES  
location/Qualifiers  
1..996  
/organism="Bos sp."  
/mol\_type="mRNA"  
/db\_xref="taxon:29061"  
/clone="C0005204p19"  
/tissue\_type="blood"  
/cell\_type="bovine monocyte"  
/clone\_lib="KN206"  
/note="Vector: pBluescriptII(SK+); Site 1: EcorV(lost); Site 2: NotI; Normalised library constructed from pooled monocytes from Bos taurus (Holstein) and Bos indicus (Sahiwal) cattle subjected to various stimuli, including infection with the protozoan parasite Theileria annulata"

## ORIGIN

Query Match 23.2%; Score 879.8; DB 1; Length 996;  
Best Local Similarity 94.9%; Pred. No. 1.3e-202;  
Matches 919; Conservative 0; Mismatches 48; Indels 1; Gaps 1;



1691 AGAAAAATGCAAAATATATTTTCCCTGGCAAAATAAAACATGCCAGTAGACAAGACGTC 1750  
Db |||||  
25 AGAAAAATGCAAAATATATTTTCCCTGGCAAAATAAAACATGCCAGTAGACAAGACGTC 84  
QY |||||  
1751 GAAACCGATGTCAGTACTGTCGATTTTCAAGATGTCAGTGTTCGATTTGTAAGAAG 1810  
Db |||||  
85 GAAACCGATGTCAGTACTGTCGATTTTCAAGATGTCAGTGTTCGATTTGTAAGAAG 144  
QY |||||  
1811 TTGTCGTCACAGATAGTCTGAAGGAGGAGAGAGTCTGCTGCTTCCAAACCAAGAGGCC 1870  
Db |||||  
145 TTGTCGTCACAGATAGTCTGAAGGAGGAGAGAGTCTGCTGCTTCCAAACCAAGAGGCC 204  
QY |||||  
1871 CATTACAAGAGAACCTTCTCAGCCCTCTCCACCTTCTCCTCCAATTCGATGATGAATG 1930  
Db |||||  
205 CATTACAGCAGGAAGCCTCTCAGCCCTCTCCACCTTCTCCTCCGATCTGTATGATGAATG 264  
QY |||||  
1931 CCCTTGTCCGAGCTTTTACAGACTCAACCCAGAGATCTTGATTTTCCAGATACTGTC 1990  
Db |||||  
265 CCCTTGTCCGAGCTTTTACAGACTCAACCCAGAGATCTTGATTTTCCAGATACTGTC 324  
QY |||||  
1991 CCACTGACCAGGCTGCTGAGGACAGATGCTGAGCATGTGCAAAATTTCTACAACCTCC 2050  
Db |||||  
325. CCACTGACCAGGCTGCTGAGGACAGATGCTGAGCATGTGCAAAATTTCTACAACCTTC 384  
QY |||||  
2051 TCACAGCCTCCATTTGATGATATCCAGAGCTGGGCAAAAGATTCGGGATTTACTGATC 2110  
Db |||||  
385 TAACAGCCTCCATTTGATGATATCCAGAGCTGGGCAAAAGATTCGGGATTTACTGATC 444  
QY |||||  
2111 TCCCAAGAGATCAGACATTTACTTTTGAATCAGCCTTTTGGAGCTTTTGGCTCTCA 2170  
Db |||||  
445 ACCCAAGAGATCAGACATTTACTTTATAGATCAGCCTTTTGGAGCTTTTGGTCTCA 504  
QY |||||  
2171 GACTTTCCATCAGTCAACACTGCTGAAGATGTTGTTCTGCAATGACTTGTC 2230  
Db |||||  
505 GACTTTCCATCAGTCAACACTGCTGAAGATGTTGTTCTGCAATGACTTGTC 564  
QY |||||  
2231 TCATCGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTGACTCTATTAAGACTTTT 2290  
Db |||||  
565 TGCACTGACTTCAGTGCCTTCGTGATTTGGGAGTGGCTGACTCCATTAAAGACTTTT 624  
QY |||||  
2291 CCTTAAATTTGAGAGCTGACCTTGATATCCAGCCTTAGCCTGCTGCTGAGCACTGA 2350  
Db |||||  
625 CTTTAAATTTGAGAGCTGACCTTGATATCCAGCCTTAGCCTGCTGCTGAGCACTGA 684  
QY |||||  
2351 GCATGATCACAAGAGACATGGTTTAAAGAACCAAGAGAGTGAAGAGCTATGCAACA 2410  
Db |||||  
685 GCATGTCACAGACGACATGGTTTAAAGAACCAAGAGAGTGGAGAGCTATGCAACA 744  
QY |||||  
2411 AGATCAAGAGCTTTTAAAGACCAACAGAGTAAGGGAAGGCTCTGGAGCCACCGAGT 2470  
Db |||||  
745 AGATCAAGAGCTTTTAAAGAGCACAGAGTAAGGGCCAGGCTTTGGAGCCACCTGAGC 804  
QY |||||  
2471 CCAAGTCTCGGTCCTGCTGAGTACAGAGATCTGAGAGATCTGACCTGGGCTCCAGCGCA 2530  
Db |||||  
805 CCAAGTCTCGGTCCTGCTGAGTACAGAGATCTGAGAGATCTGACCTGGGCTCCAGCGCA 864  
QY |||||  
2531 TCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACTTCCATCAITGACAAGCTCTTCC 2590  
Db |||||  
865 TCTTCTACCTGAGCTGGAAGACTTGGTGTCTCCACTTCCATCAITGACAAGCTCTTCC 924  
QY |||||  
2591 TGGACACCTTACTTTTCTAATCAGGAGCAG-TGGAGCAGTGAAGCTGCTCTCTCTAGC 2649  
Db |||||  
925 TGGACACCTTACTTTCTGAGCAGGAGCAGCTGAGCAGGAGCTGCTCTCTCTAGC 984  
QY |||||  
2650 ACTGCTT 2657  
Db |||||  
985 AGCTGCTT 992

RESULT 2  
BM557376 1087 bp mRNA linear EST 20-FEB-2002  
LOCUS AGENCOURT\_6561893 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5547838  
DEFINITION

5', mRNA sequence.  
BM557376  
VERSION BM557376.1 GI:18799320  
EST.  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE NIH-MSC http://mgc.nci.nih.gov/  
AUTHORS 1 (bases 1 to 1087)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
plate: LLAM12255 row: m column: 23  
High quality sequence stop: 624.  
Location/Qualifiers  
1..1087  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5547838"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

ORIGIN

Query Match 21.0%; Score 796; DB 4; Length 1087;  
Best Local Similarity 96.7%; Pred. No. 3.3e-182;  
Matches 878; Conservative 0; Mismatches 20; Indels 10; Gaps 6;  
QY 2869 GGTTCCTTCAAGAAATGCTTAACAAAGCACTTTTGGACAATGCTATCCCGAGGAGAAAAA 2928  
Db 1 GGTTCCTTCAAGAAATGCTTAACAAAGCACTTTTGGACAATGCTATCCCGAGGAGAAAAA 60  
QY 2929 AAGGATAATATACTGTTTAAACTCTTCTGGGGAATCCAATATATAGTTGCTTTGTAT 2988  
Db 61 AAGGATAATATACTGTTTAAACTCTTCTGGGGAATCCAATATATAGTTGCTTTGTAT 120  
QY 2989 TTAATAACAAGAACAGCCAGGGTGTTCGCCAGGGTAGGATGTCTTTAAAGATGGTC 3048  
Db 121 TTAATAACAAGAACAGCCAGGGTGTTCGCCAGGGTAGGATGTCTTTAAAGATGGTC 180  
QY 3049 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTTGTCACCAAGGAGGAGAACTT 3108  
Db 181 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTTGTCACCAAGGAGGAGAACTT 240  
QY 3109 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTTGTCACCAAGGAGGAGAACTT 3168  
Db 241 CTTTGAATAATGCTTCTCTGATCAAGGTAGTATGTTGTCACCAAGGAGGAGAACTT 3300  
QY 3169 AATCAGACATGGCAAAACAATAATGGCTGTTTGTTCATAAACAAAGTGAATTTTAA 3228  
Db 301 AATCAGACATGGCAAAACAATAATGGCTGTTTGTTCATAAACAAAGTGAATTTTAA 360  
QY 3229 AGTGTCTTCTTACTAAGTCTTGTGTTTAACTCTCTTATTAATGTAATGAAATAAAG 3288  
Db 361 AGTGTCTTCTTACTAAGTCTTGTGTTTAACTCTCTTATTAATGTAATGAAATAAAG 420  
QY 3289 GAGGAGTCTATGTAAGCAATGACACGTTAATATATCCCTAGCAGAGGCTGTGTTTCACTTC 3348  
Db 421 GAGGAGTCTATGTAAGCAATGACACGTTAATATCCCTAGCAGAGGCTGTGTTTCACTTC 480

QY 3349 CCTGTGATCCCTTCTGAGGTATGGCCCATCCAGACTTTTAGGCCATTTCTTGATGGAAC 3408  
 |||||  
 Db 481 CCTGTGATCCCTTCTGAGGTATGGCCCATCCAGACTTTTAGGCCATTTCTTGATGGAAC 540  
 |||||  
 QY 3409 CAGATCCCTGCCCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATATAAACTGGATTA 3468  
 |||||  
 Db 541 CAGATCCCTGCCCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATATAAACTGGATTA 600  
 |||||  
 QY 3469 CATGTAACCTGTTTGGTGTGTTCTATCAACCCACACAGATTCCTTAAACTTGTCTCAG 3528  
 |||||  
 Db 601 CATGTAACCTGTTTGGTGTGTTCTATCAACCCACACAGATTCCTTAAACTTGTCTCAG 660  
 |||||  
 QY 3529 TTATAGTAACCTGACTGTGATATTTCTTCAAGAGCGCCATAAGCTCAG-TTGAGTATTGAT 3587  
 |||||  
 Db 661 TTATAGTAACCTGACTGTGATATTTCTTCAAGAGCGCCATAAGCTCAG-TTGAGTATTGAT 720  
 |||||  
 QY 3588 CCTAGATAAGAACATCGAAATCAGCAGGAATCTGGTCATACAGGGTAAG-CACCAGGGAC 3646  
 |||||  
 Db 721 CCTAGATAAGAACCTGCGAAATCAGCAGGAATCTGGTCATACAGGGTAAGCACCAGGGAC 780  
 |||||  
 QY 3647 AATAAGGATTTTATAGATATAATTTAA--TTTTGTTATTGTTAAGGAG-ACAATTTT 3703  
 |||||  
 Db 781 AATAAGGATTTTATAGATATAATTTAAATTTTGGCTATTGGCTTAAGGAGAACATTTT 840  
 |||||  
 QY 3704 GGAGAGCAAGCAAA-----TCTTTTAAATAAGTATGATGTGTAATCTAGAAAGAT 3758  
 |||||  
 Db 841 GGAGAGCAAGCAAAATCTCTTTTAAATAAGTATGATGGGGAATACCAAAAAAC 900  
 |||||  
 QY 3759 TTAATAAA 3766  
 |||||  
 Db 901 TTTAGAA 908  
 |||||

## RESULT 3

BM549822  
 LOCUS BM549822 1032 bp mRNA linear EST 20-FEB-2002  
 DEFINITION AGENCOURT\_6544050 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5745831  
 5', mRNA sequence.  
 ACCESSION BM549822  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS 1 (bases 1 to 1032)  
 TITLE NIH-MGC http://mgc.nci.nih.gov/  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12769 row: g column: 16  
 High quality sequence stop: 651.

## FEATURES

source  
 1..1032  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5745831"  
 /tissue\_type="leukocyte"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 118"  
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV  
 (destroyed); RNA source leukocytes from anonymous pool of  
 non-activated adult donors. Library is oligo-dT primed

and directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 1.2-3.3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 027. Note:  
 this is a NIH\_MGC Library."

## ORIGIN

Query Match 20.3%; Score 769.4; DB 4; Length 1032;  
 Best Local Similarity 94.4%; Pred. No. 9.7e-176;  
 Matches 820; Conservative 0; Mismatches 46; Indels 3; Gaps 2;  
 QY 2360 CAGAAGACATGGTTAAAGAACCAAGAGAGTGAAGAGCTATCAACAGATCACAA 2419  
 Db 5 CCGGAATTCCTCGGATAAAGAACCAAGAGAGTGAAGAGCTATCAACAGATCACAA 64  
 |||||  
 QY 2420 GCAGTTTAAAGACCAACAGAGTAAGGAGCAGGCTCTGAGAGCCACCGAGTCAAGGTCC 2479  
 Db 65 GCAGTTTAAAGACCAACAGAGTAAGGAGCAGGCTCTGAGAGCCACCGAGTCAAGGTCC 124  
 |||||  
 QY 2480 TGGGTGCCCTGTAGAACTGTGAGGAAGATCTGCACCTCTGGGCTCTGAGCGCATCTTCTACC 2539  
 Db 125 TGGGTGCCCTGTAGAACTGTGAGGAAGATCTGCACCTCTGGGCTCTGAGCGCATCTTCTACC 184  
 |||||  
 QY 2540 TGAAGCTGGAAGACACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCTGGACACCC 2599  
 Db 185 TGAAGCTGGAAGACACTTGGTGTCTCCACCTTCCATCATTTGCAAGCTCTTCTGGACACCC 244  
 |||||  
 QY 2600 TACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCAGCTGCTTGC 2659  
 Db 245 TACCTTTCTAATCAGAGCAGTGGAGCAGTGCCTCTCTCTAGCAGCTGCTTGC 304  
 |||||  
 QY 2660 TAAGCAGCAAGGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTTAAGAGGAAA 2719  
 Db 305 TAAGCAGCAAGGGATAGGTTTGGAAACCTATCATTTCTGCTCTCTTAAGAGGAAA 364  
 |||||  
 QY 2720 CGAGCTCCTGTAGAAAGCAAGACTTCTTTTCTGGCTCTCTTCTCAACCTAA 2779  
 Db 365 CGAGCTCCTGTAGAAAGCAAGACTTCTTTTCTGGCTCTCTTCTCAACCTAA 434  
 |||||  
 QY 2780 AGCCAGAAAACCTTGCAGAGTATTGTGGGGTGTGTTTATATTAGGCATTTGGGGA 2839  
 Db 425 AGCCAGAAAACCTTGCAGAGTATTGTGGGGTGTGTTTATATTAGGCATTTGGGGA 484  
 |||||  
 QY 2840 TGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTAAGAAATGCTAACAAAGCACTT 2899  
 Db 485 TGGGGTGGAGGGGTTATAGTTTCATGAGGGTTTCTAAGAAATGCTAACAAAGCACTT 544  
 |||||  
 QY 2900 TTGGACAAATGCTATCCAGCAGGAAAAGGATAATATACTGTTTAAACTCTTTC 2959  
 Db 545 TTGGACAAATGCTATCCAGCAGGAAAAGGATAATATACTGTTTAAACTCTTTC 604  
 |||||  
 QY 2960 TGGGGAATCCAAATATAGTTGCTTTGTTATTTAAAAACAAGAACAGCCAGGGTTGTCG 3019  
 Db 605 TGGGGAATCCAAATATAGTTGCTTTGTTATTTAAACAAGAACAGCCAGGGTTGTCG 664  
 |||||  
 QY 3020 CAGGGTAGAGTGTCTTAAAGATTTGGTCCCTTGAATAATATGCTCTCTGATCAAAAGTTA 3079  
 Db 665 CAGGGTAGAGTGTCTTAAAGATTTGGTCCCTTGAATAATATGCTCTCTGATCAAAAGTTA 724  
 |||||  
 QY 3080 CGTATGTGGTGCAAAACAGGAGAACTTCTTTTAAATTTCTTCTCTTTATTTTAAAC 3139  
 Db 725 GTATGTGGGCAAAACAGGAGAACTTCTTTTAAATTTCTTCTCTTTATTTTAAAC 784  
 |||||  
 QY 3140 AAATGGTGAAGATGGAGGATTACCTACAAATCAGACATGGCAAAACAATAATATGGCTG-- 3197  
 Db 785 AATGGGTG-AAGATGGAGGATACCTACCAATCCAACTGGGCAACCAATAATAGGGCTGGT 843  
 |||||  
 QY 3198 TTTGCTTCCATAAAACAGTGCATTTTTT 3226  
 Db 844 TTGCTTTCCATAAAACAGGCGCAATTTTTT 872  
 |||||

RESULT 4	
BQ431513	BQ431513 819 bp mRNA linear EST 24-MAY-2002
LOCUS	AGENCOURT_7896919 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159660
DEFINITION	5', mRNA sequence.
ACCESSION	BQ431513
VERSION	BQ431513.1 GI:21170589
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 819)
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a>
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: ATCC/DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM13509 row: b column: 13 High quality sequence stop: 694.
FEATURES	Location/Qualifiers
source	1..819
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:6159660"
	/tissue_type="melanotic melanoma"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="NIH MGC 72"
	/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
ORIGIN	
Query Match	19.8%; Score 752.4; DB 5; Length 819;
Best Local Similarity	99.9%; Pred. No. 1.3e-171;
Matches 753; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	2865 TGAGGGTTTCTAAGAAATGCTAACAAAGCACTTTTGGACATGCTATCCCGACAGGAA 60
Db	1 TGAGGGTTTCTAAGAAATGCTAACAAAGCACTTTTGGACATGCTATCCCGACAGGAA 60
QY	2925 AAAAAAGGATAATTAAGTCTTTTAAACCTCTTCTGGGAATCCCAATTATAGTTCCTTT 2984
Db	61 AAAAAAGGATAATTAAGTCTTTTAAACCTCTTCTGGGAATCCCAATTATAGTTCCTTT 120
QY	2985 GTATTTAAACCAAGAACAGCCAGGGTGTTCGCCAGGGTAGGATGTGCTTTAAAGATT 3044
Db	121 GTATTTAAACCAAGAACAGCCAGGGTGTTCGCCAGGGTAGGATGTGCTTTAAAGATT 180
QY	3045 GTGTCCTTGAATAATATGCTTCCTGTATCAAGGTAGCTATGTGTCGCAACAGGCGAGAA 3104
Db	181 GTGTCCTTGAATAATATGCTTCCTGTATCAAGGTAGCTATGTGTCGCAACAGGCGAGAA 240
QY	3105 ACTTCCTTTTAAATTCCTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTACC 3164
Db	241 ACTTCCTTTTAAATTCCTCTCTCTTATTTTAAACAAATGGTGAAGATGGAGATTACC 300
QY	3165 TACAAATCAGACATGGCAACAAATAATAGCTGTTCCTTCCTCCATAACAGTGCATTTT 3224
Db	301 TACAAATCAGACATGGCAACAAATAATAGCTGTTCCTTCCTCCATAACAGTGCATTTT 360
QY	3225 TTAAGTGTCTCTTACTAAGTCTGTGTTTATTAACTCTCTCTTATTTCTATATGAATAA 3284
Db	361 TTAAGTGTCTCTTACTAAGTCTGTGTTTATTAACTCTCTCTTATTTCTATATGAATAA 420

QY	3285 AAAGGAGGAGTCATGTTAGCAAAATGACAGTGTAAATATCCCTAGCAGAGGCTGTGTTTAC 3344
Db	421 AAAGGAGGAGTCATGTTAGCAAAATGACAGTGTAAATATCCCTAGCAGAGGCTGTGTTTAC 480
QY	3345 CTTCCCTGTGATCCCTCTTCTGAGGTATGCCCATCCAGACTTTTAGGCCATTTCTTGATG 3404
Db	481 CTTCCCTGTGATCCCTCTTCTGAGGTATGCCCATCCAGACTTTTAGGCCATTTCTTGATG 540
QY	3405 GAACCATCCCTCCCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATAAACTGG 3464
Db	541 GAACCATCCCTCCCTGACTGTCCAGCTATCTGAAAGTGGATCAGATTATAAACTGG 600
QY	3465 ATTACATGTAACCTGTTTGTGTTTCTTCTATCAACCCACAGAGTTCCTTAACTTGCT 3524
Db	601 ATTACATGTAACCTGTTTGTGTTTCTTCTATCAACCCACAGAGTTCCTTAACTTGCT 660
QY	3525 TCAGTTATAGTAAGTGTGTTATATTCATTCAGAGGCGCATAGTCAGTTGAGTATTT 3584
Db	661 TCAGTTATAGTAAGTGTGTTATATTCATTCAGAGGCGCATAGTCAGTTGAGTATTT 720
QY	3585 GATCCCTAGATAAGAAATGCAATGCAATCAGCAGAA 3618
Db	721 GATCCCTAGATAAGAAATGCAATGCAATCAGCAGAA 754

RESULT 5

CK778210/c

LOCUS CK778210 852 bp mRNA linear EST 20-FEB-2004

DEFINITION 965365 MARC 3BOV Bos taurus cDNA 3', mRNA sequence.

ACCESSION CK778210

VERSION CK778210.1 GI:42730523

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 852)

AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G., Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and Keefe, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: [smith@mail.marc.usda.gov](mailto:smith@mail.marc.usda.gov)

Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim\_alt option. Vector identified with cross\_match v0.990329.

Plate: 51 row: 1 column: 19

Seq primer: GTAATACGACTCACTATAGG.

Location/Qualifiers

1..852

/organism="Bos taurus"

/mol\_type="mRNA"

/db\_xref="taxon:9913"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="MARC 3BOV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

ORIGIN

Query Match	18.4%;	Score 696.4;	DB 7;	Length 852;
Best Local Similarity	92.6%;	Pred. No. 5.5e-158;		
Matches 78;	Conservative 0;	Mismatches 56;	Indels 7;	Gaps 5;
QY	2079	CTGGGAGAGAAAGATTCCGGGATTACTGATCTCCCAAGAGATCAGACATTACTTAT	2138	
Db	849	CGGGGAGAGAAAGATTCCGGGATTACTGATCTCCCAAGAGATCAGACATTACTTAT	790	
QY	2139	TGAATCAGCCTTTTGGAGCTGTTGCTCAGACTTTTCATCAGGTCAACACTCTGA	2198	
Db	789	AGAATCAGCCTTTTGGAGCTGTTGCTCAGACTTTTCATCAGGTCAACACTCTGA	730	
QY	2199	AGATAAGTTGTGTTCTGCAATCGACTGTCCTGCATCGACTCAGTGCCTTCGTGATT	2258	
Db	729	AGATAAGTTGTGTTCTGCAATCGACTGTCCTGCATCGACTCAGTGCCTTCGTGATT	670	
QY	2259	TGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAAATTTTCAGAGCTGAACCTTGA	2318	
Db	669	TGGGAGTGGCTCGACTCCATTAAAGACTTTTCTTAAATTTTCAGAGCTGAACCTTGA	610	
QY	2319	TATCCAAAGCTTAGCTGCTGTCAGACTGAGCATGATCAGAGAGACATGGGTAAA	2378	
Db	609	TATCCAAAGCTTAGCTGCTGTCAGACTGAGCATGATCAGAGAGACATGGGTAAA	550	
QY	2379	AGAACCAAGAGAGTCCGAGAGCTATGCAACAGATCAGAGAGTTTAAAGACACCA	2438	
Db	549	AGAACCAAGAGAGTCCGAGAGCTATGCAACAGATCAGAGAGTTTAAAGACACCA	490	
QY	2439	GAGTAAGGAGACAGCTCTGGAGCCACCGAGTCCAAAGTCTTGGTGCCTTGTGAACCT	2498	
Db	489	GAGTAAGGAGACAGCTCTGGAGCCACCGAGTCCAAAGTCTTGGTGCCTTGTGAACCT	430	
QY	2499	GAGGAAGATCTGCACCTCGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAGCTTGT	2558	
Db	429	GAGGAAGATCTGCACCTCGGCTCCAGGCGATCTTCTACCTGAAGCTGGAAGCTTGT	370	
QY	2559	GTCTCCACCTCCATCATTTGCAAGCTCTTCTGGACACCTTCTTATCAGGAGC	2618	
Db	369	GTCTCCACCTCCATCATTTGCAAGCTCTTCTGGACACCTTCTTATCAGGAGC	310	
QY	2619	AG-TGGAGAGTGGAGTGGCTCTCTCTAGACCTGCTTGTCTACGAGCAAGGA-TA	2676	
Db	309	AGCTGAGCAGGAGTGGCTCTCTCTAGACCTGCTTGTCTAGTACGAGGAGG	250	
QY	2677	GGTTTGGAAACCTATCATTTCTGCTCTTCTTAAAGAG-GAAAGCAGCTCTCTGTA	2735	
Db	249	GGTCTGGACACCTTACCATTTCTGCTCTTCTTAAAGAGGAAACAAGCTCTGTTGA	190	
QY	2736	GCAAGACTTTCTTTT---TTTCTGCTCTTCTTCTTCAACCTTAAAGCCAGAACTT	2792	
Db	189	TGAAGACTTTTCTTTTCTGCTCTTCTTCTTCAACCTTAAAGCCAGAACTT	130	
QY	2793	CGACAGTATTGTGTTGGGTTGTTTATATTAGGCAATGGGGATGGGGTGGAGGG	2852	
Db	129	GCAAGTATTGTGTTGGGTTGTTTATATTAGG-TTTTGGGTTGGATGGAGGT	71	
QY	2853	GTTTATAGTATGAGGGTTTCTAAGAAATTGCTAAAGACACTTTTGGCAATGCTA	2912	
Db	70	GGTATAGTATGAGGGTTTCTAAGAAATTGCTAAAGACACTTTTGGCAATGCTA	11	
QY	2913	TCCAGCAGG 2922		
Db	10	TCCAGCAGG 1		
RESULT 6				
CB445292				
LOCUS	671 bp	mRNA	linear	EST 25-MAR-2003
DEFINITION	696544 MARC 6BOV Bos taurus cdna 5', mRNA sequence.			
ACCESSION	CB445292			
VERSION	CB445292.1			
KEYWORDS	EST.			
SOURCE	Bos taurus (cow)			
ORGANISM	Bos taurus			
REFERENCE	1 (bases 1 to 671)			
AUTHORS	Smith, T.P.L., Roberts, A.J., Echternkamp, S.E., Chitko-McKown, C.G., Wray, J.E. and Keele, J.W.			
TITLE	A second set of bovine ESTs from pooled-tissue normalized libraries			
JOURNAL	Unpublished (2003)			
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329. Plate: FQ18053 row: 1 column: 18 Seq primer: GTAATACGACTCACTATAGGG.			
FEATURES	Location/Qualifiers			
source	1..671			
	/organism="Bos taurus"			
	/mol_type="mRNA"			
	/db_xref="taxon:9913"			
	/tissue_type="pooled"			
	/lab_host="DH10B"			
	/clone_lib="MARC 6BOV"			
	/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."			
ORIGIN				
Query Match	16.5%;	Score 627.8;	DB 6;	Length 671;
Best Local Similarity	96.0%;	Pred. No. 2.6e-141;		
Matches 644;	Conservative 0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	1755	CCGATGTCAGTACTGTCGATTTTCAGAGTGTCTCAGTGTGGAATGGTAAAGAGCTTGT	1814	
Db	1	CCGATGTCAGTACTGTCGATTTTCAGAGTGTCTCAGTGTGGAATGGTAAAGAGCTTGT	60	
QY	1815	CCGTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCAAACCAAGAGCCCAT	1874	
Db	61	CCGTACAGATAGTCTGAAAGGAGGAGAGTGTCTGCTTCCAAACCAAGAGCCCAT	120	
QY	1875	ACAAAGAGAACTTCTCAGCCCTCTCCACCTTCTCTCCAACTGCATCATGAATGCCCT	1934	
Db	121	ACAGCAGGAGGCTCTCAGCCCTCTCCACCTTCTCTCCAACTGCATCATGAATGCCCT	180	
QY	1935	TGTCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTATTCAGATCTGTCCCAC	1994	
Db	181	TGTCGAGCTTTAACAGACTCAACACCCAGAGATCTTGATTATTCAGATCTGTCCCAC	240	
QY	1995	TGACAGAGCTGTGAGGACAGATGTCGAGATGTGCAACAAATTTCTACAACTCTCTGAC	2054	
Db	241	TGACAGAGCTGTGAGGACAGATGTCGAGATGTGCAACAAATTTCTACAACTCTCTGAC	300	
QY	2055	AGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCGGGATTTTACTGATCTCC	2114	
Db	301	AGCCTCCATTGATGATCCAGAGCTGGGAGAGAAAGATTCGGGATTTTACTGATCTCC	360	
QY	2115	CAAGAAGATCAGACATTACTTATTAAGATCAGCCCTTTTGGAGCTGTTTGTCTCAGACT	2174	
Db	361	CAAGAAGATCAGACATTACTTATTAAGATCAGCCCTTTTGGAGCTGTTTGTCTCAGACT	420	
QY	2175	TTCCATCAGGTCAACACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCA	2234	
Db	421	TTCCATCAGGTCAACACACTGCTGAAGATAAGTTTGTGTTCTGCAATGGACTTGTCTGCA	480	
QY	2235	TCGACTTCAGTCCCTTCGTGGATTTCGGGAGTGGCTCCGACTCTATTAAAGACTTTTCCCT	2294	

481	TCGAC	TTCAG	TGCCTTCGT	GGATTCGTG	GAATTTGGGGAGTGGCTCG	CACTCCATTAAGAC	CTTTTCCTT	540		
	Db									
2295	AAATTTG	CAGAGCCTG	AAACCTTCG	ATATCA	AGCCTTTAG	CCTCCTGTG	CAGAC	CTGAGCAT	2354	
	Qy									
541	AAGTTTG	CAGAGCCTG	AACTTCG	ATATCA	AGCCTTTAG	CCTCCTGTG	CAGAC	CTGAGCAT	600	
	Db									
2355	GATCACA	GAAGA	CATGGG	TTAAAG	AACCAAG	AGAGCTTCG	AAAGAGCT	TATGCA	CAAGAT	2414
	Qy									
601	GGTCACA	GAACG	CATGGG	TTAAAG	AACCAAG	AGAGCTTCG	AAAGAGCT	TATGCA	CAAGAT	660
	Db									
2415	CACAAG	CAGTT	2425							
	Qy									
661	CACAAG	CAGCT	671							
	Db									

RESULT 7	CD521819	757 bp	mRNA	linear	EST 06-JUN-2003
LOCUS	AGENCE0014360118	NIH MGC 191	Homo sapiens	CDNA clone	
DEFINITION	IMAGE:30409856	5',	mRNA	sequence.	

ESL.  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
1 (bases 1 to 757).  
AUTHORS  
NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Daniela S. Gerhard, Ph.D.

## ORIGIN

Query Match 16.3%; Score 617; DB 6; Length 757;  
Best Local Similarity 98.9%; Pred. No. 1.2e-138;  
Matches 632; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

2866	QY	GAGGGTTTTCTAAGAAATGCTTAA	CAAGCACTTTTGGCAATGCTATCC	CAGCAGAAA	2922
1	Db	GGGGTTTTCTAAGAAATGCTTAA	CAAGCACTTTTGGCAATGCTATCC	CAGCAGAAA	60
2926	QY	AAAAAGGATAATATACTGTTTTTAA	AACTCTTCTGGGGAATCCAATATAGT	TGCTTTG	2985
61	Db	AAAAAGGATAATATACTGTTTTTAA	AACTCTTCTGGGGAATCCAATATAGT	TGCTTTG	120
2986	QY	TATTTAAAAACAAGAACAGCAGGGT	CTGCCCAGGGTAGGATGTCTTAA	AGATTG	3045
121	Db	TATTTAAAAACAAGAACAGCAGGGT	CTGCCCAGGGTAGGATGTCTTAA	AGATTG	180
3046	QY	GTCCCTTGAAAATATGCTTCTGTAT	CAAGGTACGTATGCGTCAAAACAAGC	CAGAAA	3105
181	Db	GTCCCTTGAAAATATGCTTCTGTAT	CAAGGTACGTATGCGTCAAAACAAGC	CAGAAA	240
3106	QY	CTTCCCTTTAAATTCCTCTCTCTT	ATTTTAAACAAATGGTGAAAGTGA	GAGATTACCT	3165
241	Db	CTTCCCTTTAAATTCCTCTCTCTT	ATTTTAAACAAATGGTGAAAGTGA	GAGATTACCT	300
3166	QY	ACAAATCAGACATGGCAAAACAAT	ATATGCTGTTCTCCATPAAACAAGT	GCAATTTTT	3225
301	Db	ACAAATCAGACATGGCAAAACAAT	ATATGCTGTTCTCCATPAAACAAGT	GCAATTTTT	360
3226	QY	TAAAGTGCTGTCTTACTAAGTCTT	GTGTTATTAATCTCTCCTTTATCT	ATATGGAATAAA	3285
361	Db	TAAAGTGCTGTCTTACTAAGTCTT	GTGTTATTAATCTCTCCTTTATCT	ATATGGAATAAA	420
3286	QY	AAGGNGCAGTCATGTTAGCAATG	ACACGTTAATATCCCTAGCAGAGG	CTGTGTTCAAC	3345
421	Db	AAGGAGCAGTCATGTTAGCAATG	ACACGTTAATATCCCTAGCAGAGG	CTGTGTTCAAC	480
3346	QY	TTCCCTGTCCGATCCCTTCTGAGT	ATGGCCCATCCCAAGACTTTTAG	GCCATCTTGATGG	3405
481	Db	TTCCCTGTCCGATCCCTTCTGAGT	ATGGCCCATCCCAAGACTTTTAG	GCCATCTTGATGG	540
3406	QY	AACCAAGATCCCTGCCCTGACTGT	CCAGCTATCCTGAAAGTGGATC	AGATTATAAACTGGA	3465
541	Db	AACCAAGATCCCTGCCCTGACTGT	CCAGCTATCCTGAAAGTGGATC	AGATTATAAACTGGA	600
3466	QY	TTACATGTAACATGTTTTT	-GGTGTGTTCTATCAACCCC	3502	
601	Db	TTACATGTAACATGTTTTT	GGTGTGTTCTATCAACCCC	639	

## RESULTS

CB465379				EST. 26-MAR-2003
LOCUS	679 bp	mRNA	linear	
SEQUENCE	Bovine	cDNA	5'	mRNA sequence.

DEFINITION  
ACCESS TO

VERSION

## KEYWORDS

**SOURCE**  
**ORIGIN**

## ORGANIC

## REFERENCE

## AUTHOR

TITLE

JOURNAL

**COMMENT**

[illegible]

Db 481 GATTACTGATCCCCAAAGAGATCAGACATTACTTATAGATCAGCCCTTTGGAGC 540  
 Qy 2159 TGTGTGCTCAGACTTTCATCAGGTCAACACAGCTGCTGAAGATAAGTTGTGTCTGCA 2218  
 Db 541 TGTGTGCTCAGACTTTCATCAGGTCAACACAGCTGCTGAAGATAAGTTGTGTCTGCA 600  
 Qy 2219 ATGGACTGCTGCTGCATCAGCTTCACTGCTTCTGCT 2253  
 Db 601 ATGGACTGCTGCTGCATCAGCTTCACTGCTTCTGCT 635

RESULT 10  
 BU436288 940 bp mRNA linear EST 29-NOV-2002  
 LOCUS 603208116F1 CSEQRBN11 Gallus gallus cDNA clone CHEST19303 5', mRNA  
 DEFINITION  
 ACCESSION BU436288  
 VERSION BU436288.1 GI:25925599  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus

REFERENCE  
 AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 TITLE A comprehensive collection of Chicken cDNAs  
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)  
 MEDLINE 22335534  
 PUBMED 12445392

COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers

FEATURES  
 source  
 1..940  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer and broiler"  
 /db\_xref="taxon:9031"  
 /clone="ChEST18303"  
 /sex="Male and female"  
 /tissue\_type="muscle"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="CSEQRBN11"  
 /note="Vector: pBluescript II KS(+); Site 1: EcoRI;  
 Site 2: NotI; This normalized library was constructed from  
 1 million independent clones. cDNA synthesis was initiated  
 using an oligo(dT) primer, using methylated C in the first  
 strand synthesis reaction. Following this first strand  
 reaction, double-stranded cDNA was blunted, ligated to  
 NotI adapters, digested with EcoRI, size-selected, and  
 cloned into the NotI and EcoRI compatible sites of a  
 custom modified MCS of the pBluescript (KS+) vector. The  
 library was normalized in 2 rounds using conditions  
 adapted from Soares et al., PNAS (1994) 91: 9228-9232 and  
 Bonaldo et al., Genome Research 6 (1996): 791, except that  
 a significantly longer reannealing hybridization was  
 used."

## ORIGIN

Query Match 15.4%; Score 583; DB 5; Length 940;  
 Best Local Similarity 81.8%; Pred. No. 2.3e-130;  
 Matches 703; Conservative 0; Mismatches 140; Indels 16; Gaps 2;  
 Qy 1767 CTGTGATTTTCAAGAGTCTCTCAGTGTGGTAAGTGAAGAGTGTCCGTACAGATAG 1826

Db 2 CTGCGGTTTTCAAGAGTGTCTCAGCGTGGCATGGTGAAGAAGTTGTCCGCAAGGACAG 61  
 Qy 1827 TCTGAAGAGGAGAGAGGTCGTCTGCTTCCAAACCAAGAGAGCCATTACAAACAGGAACC 1886  
 Db 62 CTTGAAGAGGAGAGAGGTCGTCTGCTTCCAAACCAAGAGAGCCCTCGCAGCAGGAACC 121  
 Qy 1887 TTCTCAGCCCTCTCCACCTTCTCTCCCAATCTCGATGATGAATGCCCTTGTCCGAGCTTT 1946  
 Db 122 TTGCGAGCCCTCCCGGCTTCTCTCTCCCATCAGCATGATGAACGCCCTCGTAGAGCTTT 181  
 Qy 1947 AACAGACTCAACCCAGAGATCTTGATTATTCAGATATCTCCCACTCACCAGGCTGC 2006  
 Db 182 AACCGACTCCAGCCCGAGGAGCTCGACTATTACGATATCTGTTCCACCGATCAGGCTGC 241  
 Qy 2007 TGCAGGCACAGATCTCAGCATGTGCAACAATTTCTACAACCTCTCTGACAGCCTCCATTGA 2066  
 Db 242 TGCAGGCACAGATGCAACAATGTACAACAGTTTCTATAATCTTCTGACTGCTCCATTGA 301  
 Qy 2067 TGTATCCAGAGCTGGGCGAGAAAGATTCCGGGATTTACTGATCTCCCAAGAGATCA 2126  
 Db 302 CATATCCAGAGCTGGGCGAGAAATTTCCAGGATTTACTGACCTCCCGAAGAGATCA 361  
 Qy 2127 GACATTACTTATTCGAATCAGCTTTTGGAGCTGTTTGTCTCAGACTTTCCATCAGGTC 2186  
 Db 362 GACATTACTTATTCGAATCAGCTTTTGGAGCTGTTTGTCTCAGACTTTCCATCAGGTC 421  
 Qy 2187 AAACACTGCTGAAAGATTTGTCTGCAATGGAATTTCTGATGATCTCTGATGATGATG 2246  
 Db 422 TGATACTGCTGAGGATAAGTTTGTATTTCTGCAATGGACTTGTGCTTCATAGACTTCAGT 481  
 Qy 2247 CTTCTGTTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAATTTGCAAG 2306  
 Db 482 CTTCTGTTGATTTGGGAGTGGCTCGACTCTATTAAAGACTTTTCTTAATTTGCAAG 541  
 Qy 2307 CTTGAACCTTGCATATCCAAAGCTTTAGCTGCTCTCAGACTGATGATGATGATGATG 2366  
 Db 542 CTTGAACCTTGCATATCCAAAGCTTTAGCAAGTTTATCAGCTCTAATCTATGATGATG 601  
 Qy 2367 ACATGGGTTAAAGAACCAAGAGAGTGAAGAGCTATGCAAGATGCAAGATGCAAGCAGTTT 2426  
 Db 602 ACATGGCTTAAAGAACCAAGAGAGTGAAGAGCTATGCAAGATGCAAGATGCAAGCAGTTT 661  
 Qy 2427 AAAGACCA-----CCAGAGTAAGGAGCAGAGCTCTGGAGCCACCGAGTC 2471  
 Db 662 GAAGATCATTAACTTTCAGTTGCCAAACAAAGGACAGCGCTCGAGTCGAGAGGCC 721  
 Qy 2472 CAAGGTCTCTGGTGCCTTGGTAGAACTCAGAGAAATCTGCACCTTGGGCTCCAGCGCAT 2531  
 Db 722 TAAGGTACTGGGTGTTCTTGTCTG-CTTCTGCACTCTGGGACTGCAAGCGCAT 780  
 Qy 2532 CTTTACCTGAAGCTGGAAGACTTGGTGTCTCCACCTTCCATCATTTGACAAAGCTCTTCT 2591  
 Db 781 CTTTACCTGAAGCTGGAAGACTTGGTGTGGGTGAGCCCTTCCAAATATCGAAGCTCTTCT 840  
 Qy 2592 GCACACCTTACCTTTCTTAA 2610  
 Db 841 GGACAACTTTCCTTCTGA 859

## RESULT 11

CD638747

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

CD638747 635 bp mRNA linear EST 17-JUN-2003  
 AGENCOURT\_14534944 NIH MGC\_191 Homo sapiens cDNA clone  
 IMAGE:30417493 5', mRNA sequence.

CD638747  
 CD638747.1 GI:31804872  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 635)





RESULT 13	CB447773	689 bp	linear	EST 26-MAR-2003
LOCUS	701798 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.			
DEFINITION	CB447773			
ACCESSION	CB447773.1	GI:29254155		
VERSION	EST.			
KEYWORDS	Bos taurus (cow)			
ORGANISM	Bos taurus			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
	Bovinae; Bos.			
REFERENCE	1 (bases 1 to 689)			
AUTHORS	Smith, T.P.L., Roberts, A.J., Ehternkamp, S.E., Chitko-McKown, C.G., Wray, J.E. and Keele, J.W.			
TITLE	A second set of bovine ESTs from pooled-tissue normalized libraries			
COMMENT	Unpublished (2003)			
	Contact: Smith TPL			
	USDA, ARS, US Meat Animal Research Center			
	PO Box 166, Clay Center, NE 68933-0166, USA			
	Tel: 402 762 4366			
	Fax: 402 762 4390			
	Email: smith@email.marc.usda.gov			
	Single pass sequencing. Bases called with phred v0.020425.c and			
	trimmed with the aid of the trim_alt option. Vector identified with			
	cross match v0.990329.			
	Plate: PQY8056 row: B column: 15			
FEATURES	Seq primer: TAGAAGGCACACTCGAGG.			
source	Location/Qualifiers			
	1..689			

RESULT 14  
CK635592  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 802)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Dr. James Lin University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)  
The following repetitive elements were found in this cDNA  
sequence: 86-116, >GC-rich#low\_complexity  
Seq primer: pyx-5.  
Location/Qualifiers  
1..802  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30643330"  
/tissue\_type="Upper Head"  
/dev\_stage="9.5-10.5 dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH-BMAP-RN0"  
/notes="Organ: Head; Vector: pYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CGACTGAT. This library was created for the University  
Iowa Brain Anatomy Project (BMAP): 'Gene discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH)."  
ORIGIN  
Query Match 14.6%; Score 555.4; DB 7; Length 802;  
Best Local Similarity 85.8%; Pred. No. 1.2e-123;  
Matches 688; Conservative 0; Mismatches 104; Indels 12; Gaps 6;  
QY 148 ACACAGCGCGCACAGCGTCCGCGACACACTCTCGCTCTCCCGCGGCTACACCCCT 207  
DB 2 ACAGCGCGCTACAGCTCCGCGACACACTCTCTCTCCCGCGGCTACACCCCT 61  
QY 208 CTTGCCCTGAGCCCTTCCCGGT-----GCAGCGCGCGCGCGAGCTGAGCGCCCTCCCG 262  
DB 62 CTCACCGCGTCCCTCGCAGTGTCTCGCGCGCGCGCGCGCGCGCGCGCCCTCCAG 121  
QY 263 GGTCTACTTGGACCTGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 322  
DB 122 GGTCTACTTGGACCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181  
QY 323 ATCTCTCCCTCTGTCACAGCCCCAACGAGGAGCGCGCGGACCTCTCGGCTGTCTCT 382  
DB 182 ATCTTACCCCTGCTGCGAGCCGAGACTGGATGCTCTGGAACTCTCGCGCGCGCTCT 241  
QY 383 CCCATGAGTGGGATCGCAGCATCCCCCACCAGCGCTCACCGCCCTCGGAGCGCGCTGG 442  
DB 242 CCCATGAGTGGGATCGCAGCATCCCCCGCGCGCGCGCTCACCGCCCTCGGAGCGCGCTGG 301  
QY 443 GCTTGTACCGCGCGCCCTTCGGGGACAGCAGCTGTGACTCTCCCCCGCGAGTGTTCG 502

Db 302 GTTGTGACCCGAGCCCTTCGGGACACAGCTGTGACTCTCCCCCAATCCAGATTTCG 361  
QY 503 GGACAGCTCTCTAGAACTCTGCTCTAAAGACGGAACCGGCACAGCACTCAAGGCCACTG 562  
Db 362 GGTGCGCTCTCTAGAACTCTGCTCTAAAGACGGAACCGGCACAGCACTCAAGGCCACTG 421  
QY 563 CGGAGAGGGGACCGCGGCAAGCGCGGCGCTGAGCTGGAGCCCTTAGGGGTGCGGGGCA 622  
Db 422 CGGAGAGGGGACCGCGGCAAGCGCGGCGCTGAGCTGGAGCCCTTAGGGGTGCGGGGCA 481  
QY 623 GCA---CTGCGCGCGCTTCGCTCGCGGACGCTCGCTCTCTACACTCTCAGCTCCG 679  
Db 482 GCACAGCGGCTGCTGCTTTCGCTTATCCGAGCTGCCCGCTCTCTACACTCTCAGCTCCG 541  
QY 680 CTTGAGAGAGCCCCCAGCGCCCACTTACGCGGCGAAGATACCTTCAGATATGCCCTGCG 739  
Db 542 CTTGAGAGAG-CCCCAGCGCCCACTTACGCGGCGAAGATACCTTCAGATATGCCCTGCG 600  
QY 740 TCAAGCCCAATATAGCCCTTCCTCCCTCAGGTTTCAGTTATGGGGGCGACATACAGCT 799  
Db 601 TGCAGCGCCAGTATAGCCCTTCACCTCCGGGGTCCACTTACGCCACGCACTTATGGCT 660  
QY 800 CGGAATACACCGGAGATCATGACCCGAGCTTACACCAAGCTGACCATGGACCTTGGCA 859  
Db 661 CGGAATA-ACCACAGAAATCATGAA-CCGACTTACACCAAGCTGACCATGGACCTTGGTA 718  
QY 860 GCACCTGAGATCAGGCTCAGCGCACCGACCTCCCTGCGCGAGCATCAGTACCTTCGTGGAGG 919  
Db 719 GCACGGGATCATGNNACCGCCACTTACATCCCTGCCAGCTTCAGTACCTTCATGAGG 778  
QY 920 GCTACTCGAGAACTACGAATCA 943  
Db 779 GCTAC-CCAGCAGCTCGCAACTCA 801  
CB445564 644 bp mRNA linear EST 26-MAR-2003  
696928 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.  
CB445564  
CB445564.1 GI:29251946  
EST.  
Bos taurus (cow)  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
1 (bases 1 to 644)  
Smith,T.P.L., Roberts,A.J., Ehternkamp,S.E., Chitko-McKown,C.G.,  
Wray,J.E. and Keele,J.W.  
A second set of bovine ESTs from pooled-tissue normalized libraries  
Unpublished (2003)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: [smith@email.marc.usda.gov](mailto:smith@email.marc.usda.gov)  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified with  
cross\_match v0.990329.  
Plate: PQY8053 row: I column: 18  
Seq primer: TAGAAGGCACAGTCGAGG.  
FEATURES  
Location/Qualifiers  
1..644  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 6BOV"  
/note="Vector: pCDNA3.1; Site\_1: EcoRI; Site\_2: NotI;



GenCore version 5.1.6  
Copyright (C) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 17:36:21 ; Search time 170 Seconds  
(without alignments)  
1424.188 Million cell updates/sec

Title: US-10-608-863-2  
Perfect score: 3337  
Sequence: 1 MFCVQAQSPSPGSSAAQ.....EDLVSPPSIIDKFLDTLPF 626

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3337	100.0	626	8	AD113008 Human NOR
2	3333	99.9	626	7	ADb45995 Human Pro
3	3333	99.9	626	7	ADf28879 Human nor
4	3333	99.9	626	8	ADp05709 Human nuc
5	3316	99.4	626	8	ADL57171 Human NOV
6	3301	98.9	626	2	AAW16398 Human neu
7	3299.5	98.9	625	5	AAU96995 Human nuc
8	3299.5	98.9	625	7	ADf28878 Human nor
9	3133.5	93.9	643	7	ADf28862 Pig nor-1
10	3123	93.6	587	8	ADL83262 Human PRO
11	3123	93.6	587	8	ADr14063 Human NF-
12	3074	92.1	628	2	AAR92057 Apoptotic
13	3074	92.1	628	5	ABb98438 Murine Ne
14	3074	92.1	628	7	ADd45993 Rat Prote
15	3074	92.1	628	7	ADf28880 Rat nor-1
16	3054.5	91.5	627	5	ABb98437 Murine Ne
17	2099.5	62.9	446	7	ADf28881 Pig nor-1
18	2036.5	61.0	429	8	ADp05711 Mouse nuc
19	1751	52.5	598	2	AAr48631 Sequence
20	1751	52.5	598	4	AAU09156 Human NOT
21	1751	52.5	598	4	AAb35156 Human nuc
22	1751	52.5	598	6	ABr42448 Human Nur
23	1751	52.5	598	7	ADb84066 Human NUR
24	1751	52.5	598	7	ADr384062 Human NUR
25	1751	52.5	598	7	ADb84067 Human NUR

26	1751	52.5	598	7	ADb84068 Human NUR
27	1751	52.5	598	7	ADb83989 Human NUR
28	1751	52.5	598	7	ADb84070 Human NUR
29	1751	52.5	598	7	ADe56986 Human Pro
30	1751	52.5	598	7	ADe25771 Human pro
31	1751	52.5	598	8	ADH75170 Human tra
32	1751	52.5	598	8	ADp05705 Human nuc
33	1751	52.5	598	8	ADp23740 PRO polyp
34	1746	52.3	598	5	AAU96994 Mouse nuc
35	1746	52.3	598	7	ADb84074 Murine NU
36	1746	52.3	598	7	ADb84065 Murine NU
37	1746	52.3	598	7	ADb84064 Murine NU
38	1746	52.3	598	7	ADb84061 Murine NU
39	1746	52.3	598	8	ADL70235 Mouse Nur
40	1746	52.3	598	8	ADp05707 Mouse nuc
41	1736	52.0	598	7	ADb84069 Rat NURR1
42	1736	52.0	598	7	ADe56984 Rat Prote
43	1732	51.9	607	4	AAU09018 Mouse Flg
44	1684.5	50.5	597	6	ABr42447 Rat Nurrl
45	1684.5	50.5	597	7	ADf28899 Rat nur77

## ALIGNMENTS

RESULT 1  
AD113008  
ID AD113008 standard; protein; 626 AA.

XX AC AD113008;  
XX DT 22-APR-2004 (first entry)  
XX DE Human NOR-1 (MINOR) protein sequence SeqID 2.  
XX KW human; receptor; allergic disease; NOR-1; MINOR; eosinophil;  
XX KW atopic dermatitis; antiallergic; antiinflammatory; dermatological.  
XX OS Homo sapiens.  
XX PN WO2004003198-A1.  
XX PD 08-JAN-2004.  
XX PF 27-JUN-2003; 2003WO-JP008199.  
XX PR 27-JUN-2002; 2002JP-00188490.

XX (GENO-) GENOX RES INC.  
XX (NIGE-) JAPAN GEN AGENCY NATION.  
XX Hashida R, Kagaya S, Yayoi Y, Sugita Y, Saito H;  
XX WPI: 2004-083057/08.  
XX N-PSDB; AD113007.  
XX Examining allergic diseases e.g. atopic dermatitis by differential display based on gene expression of NOR-1 receptor protein, also applicable in screening compounds for treatment of allergic diseases.  
XX Example 1; SEQ ID NO 2; 155pp; Japanese.

XX This invention relates to a novel method for examining allergic diseases that comprises comparing the expression levels of a gene encoding the NOR-1 receptor protein between patients and healthy individuals.  
XX Specifically, the NOR-1 gene, also referred to as MINOR, is expressed in the specialist white blood cells known as eosinophils and is involved in mediating an allergic reaction. The present invention describes a differential display method that can identify the expression level of this gene in order to identify its usefulness in diagnosing allergic diseases such as atopic dermatitis. Furthermore, compositions can also be used to screen compounds for the treatment of allergic diseases.  
XX Accordingly, they exhibit various activities including antiallergic,



QY 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360  
DB 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360  
QY 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420  
DB 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420  
QY 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAFLEFVL 480  
DB 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAFLEFVL 480  
QY 481 RLSIRSNATDAEDKFVFCNGVLVHLRQLRGFGWLSIKDFSLNLSLNLDIQAALCSAL 540  
DB 481 RLSIRSNATDAEDKFVFCNGVLVHLRQLRGFGWLSIKDFSLNLSLNLDIQAALCSAL 540  
QY 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKTCTIGLOR 600  
DB 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKTCTIGLOR 600  
QY 601 IFYKLEDLVSPPSIIDKFLDLPF 626  
DB 601 IFYKLEDLVSPPSIIDKFLDLPF 626

RESULT 3

ADP28879

ID ADF28879 standard; protein; 626 AA.

AC ADF28879;

XX 12-FEB-2004 (first entry)

XX Human nor-1 nuclear receptor polypeptide Q92570.

XX Human; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;

XX cytostatic; haematopoietic cell.

XX Homo sapiens.

XX WO2003088812-A2.

XX 30-OCT-2003.

XX 15-APR-2003; 2003WO-US011804.

XX 17-APR-2002; 2002US-0373238P.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

XX (UNIW ) UNIV WASHINGTON.

XX Mulligan SE, Conneely OM, Milbrandt J;

XX WPI: 2003-854017/79.

XX GENBANK; Q92570.

XX Inhibiting proliferation of a hematopoietic cell, useful for treating or

XX preventing leukemia, comprises modulating the level or activity of nor-1

XX and/or nur77 nuclear receptor.

XX Disclosure; SEQ ID NO 14; 101pp; English.

XX The present sequence is that of human nor-1 nuclear receptor polypeptide

XX Q92570. Nor-1 has been identified as a molecular target for therapeutic

XX intervention in the treatment of myeloid leukaemia. The invention

XX provides methods of inhibiting the proliferation of a haematopoietic stem

XX cell or a haematopoietic myeloid cell by modulating the level of a nor-1

XX and/or nur77 nuclear receptor. This involves increasing the level of the

XX receptor polypeptide or polynucleotide, e.g. by administration of a

XX vector comprising the polynucleotide. A claimed method of treating

XX leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a

XX haematopoietic stem cell or myeloid cell. Also claimed are methods of

CC identifying an upregulator of expression of nor-1 and/or nur77, of  
CC identifying a compound for the treatment of leukaemia, and of screening  
CC for a compound for treatment of leukaemia, and a mouse model for  
CC leukaemia comprising a mouse having defective nor-1 and nur77 nucleic  
CC acid sequences.

XX SQ Sequence 626 AA;

Query Match 99.9%; Score 3333; DB 7; Length 626;

Best Local Similarity 99.8%; Pred. No. 3.1e-222;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPCVQAQYSPSPGSSYAAQTYSSYTTIEMNPDYTKLTMDLGSITEITATATSLPSIST 60

DB 1 MPCVQAQYSPSPGSSYAAQTYSSYTTIEMNPDYTKLTMDLGSITEITATATSLPSIST 60

QY 61 FVEGYSSNYELKPSCVYQMRPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHH 120

DB 61 FVEGYSSNYELKPSCVYQMRPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHH 120

QY 121 SSPDEVLPSTSMYFKQSPPTPTTAPFPQAGALWDEALPSAPGCIAPGLLDPKMKAV 180

DB 121 SSPDEVLPSTSMYFKQSPPTPTTAPFPQAGALWDEALPSAPGCIAPGLLDPKMKAV 180

QY 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALBS 240

DB 181 PTVAGARFPLFHFKPSPPHPPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALBS 240

QY 241 HPYGLPLAKRAAPLAPPLGLTPTSTASSLLGESPSLPSPPSRSSSSSGGTCAVCGDAA 300

DB 241 HPYGLPLAKRAAPLAPPLGLTPTSTASSLLGESPSLPSPPSRSSSSSGGTCAVCGDAA 300

QY 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360

DB 301 COHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRQYCRFQKCLSVGMVKE 360

QY 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420

DB 361 VVRTDSLKGRRLPSKPSPLQOEPSQSPSPPICMNVALVRALTDSTPRDLDSRYC 420

QY 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAFLEFVL 480

DB 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAFLEFVL 480

QY 481 RLSIRSNATDAEDKFVFCNGVLVHLRQLRGFGWLSIKDFSLNLSLNLDIQAALCSAL 540

DB 481 RLSIRSNATDAEDKFVFCNGVLVHLRQLRGFGWLSIKDFSLNLSLNLDIQAALCSAL 540

QY 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKTCTIGLOR 600

DB 541 SMITERHGLKEPKRVEELCNKITSSLDKHSQKQALEPTESKVLGALVELRKTCTIGLOR 600

QY 601 IFYKLEDLVSPPSIIDKFLDLPF 626

DB 601 IFYKLEDLVSPPSIIDKFLDLPF 626

RESULT 4

ADP05709

ID ADP05709 standard; protein; 626 AA.

XX AC ADP05709;

XX DT 26-AUG-2004 (first entry)

XX DE Human nuclear receptor protein SeqID83.

XX KW disease risk; disorder risk; mutation; polymorphism;

XX KW nuclear receptor protein; antibacterial; antithyroid; cardiovascular-Gen;

XX KW cytostatic; dermatological; eating-Disorders-Gen; gastrointestinal-Gen;

XX KW gynaecological; hepatotropic; immunosuppressive; muscular-Gen;

XX KW nephrotropic; osteopathic; virucide; adrenal gland; colon;

XX KW cardiovascular; intestine; kidney; liver; lung; muscular; ovary; blood;





PI Zhong M, Guo X, Anderson DW, Ort T, Padigaru M, Rieger DK;  
XX WPI; 2004-315567/29.  
DR N-PSDB; ADU571170.

PT New isolated NOVX polypeptides and polynucleotides, useful for  
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.  
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,  
PT asthma, or infections.

XX Claim 1; SEQ ID NO 116; 214bp; English.

XX The invention relates to a novel isolated polypeptide (NOVX) comprising a  
CC mature form of any of the 37 amino acid sequences fully defined in the  
CC specification. A polypeptide of the invention has antidiabetic,  
CC anorectic, cardiac, hypotensive, antiarteriosclerotic, anorectic,  
CC viricide, antibacterial, fungicide, protozoacide, nootropic,  
CC neuroprotective, antiParkinsonian, anticonvulsant, osteopathic,  
CC antiarthritic, antiinflammatory, dermatological, antiasmatic, and  
CC antilipemic activity. A polynucleotide of the invention may have a use  
CC in gene therapy. The polypeptides, nucleic acid molecules and antibodies  
CC are useful in the manufacture of a medicament for treating a syndrome  
CC associated with a human disease, preferably a NOVX-associated disorder.  
CC The nucleic acid molecules, polypeptides and antibodies are useful for  
CC treating, preventing or diagnosing diseases such as metabolic disorders,  
CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,  
CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases  
CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's  
CC disease, Parkinson's disease, epilepsy, immune disorders  
CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,  
CC asthma, and various dyslipidaemias. The nucleic acids and polypeptides  
CC may also be used as targets for the identification of small molecules  
CC that modulate or inhibit e.g. neurogenesis, cell differentiation, cell  
CC proliferation, haematopoiesis, wound healing and angiogenesis, in gene  
CC therapy, in generation of antibodies that bind immunospecifically to NOVX  
CC substances for use in therapeutic or diagnostic methods. The nucleic  
CC acids are further used as hybridisation probes, in chromosome mapping,  
CC tissue typing, preventive medicine, and pharmacogenomics. The NOVX  
CC polypeptides of the invention show homology to certain known human  
CC proteins: NOV1a-1t show homology to fibroblast growth factor receptor 4  
CC (FGFR4); NOV2a shows homology to complement factor I precursor; NOV3a  
CC shows homology to matrix metalloproteinase-15 precursor; NOV4a shows  
CC homology to MDC3; NOV5a-5c show homology to T-lymphocyte surface antigen  
CC Ly-9 precursor; NOV6a-6m show homology to fibroblast growth factor-21  
CC (FGF-21); NOV7a-7c show homology to alpha-2 macroglobulin-like  
CC polypeptide variant; NOV8a-8g show homology to antileukoproteinase 1  
CC precursor; NOV9a-9i show homology to LIV-1 protein; NOV10a shows homology  
CC to nuclear hormone receptor NOR-1; NOV11a-11j show homology to  
CC transmembrane protein-like; NOV12a-12c show homology to beta-neoendorphin  
CC -dynorphin precursor. The present sequence represents a NOVX polypeptide  
CC of the invention.

XX Sequence 626 AA;

Query Match 99.4%; Score 3316; DB 8; Length 626;  
Best Local Similarity 99.4%; Pred. NO. 4.6e-221;  
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MPCVQAYSPSPGSSYAAQYSSSEYTTIMNPDYTKLTWDLGSTEITATATSLPSIST 60  
DB 1 MPCVQAYSPSPGSSYAAQYSSSEYTTIMNPDYTKLTWDLGSTEITATATSLPSIST 60  
QY 61 FVEGYSSNVELKPSQYVQORPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 120  
DB 61 FVEGYSSNVELKPSQYVQORPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 120  
QY 121 SSPDEVLSTSMYFKQSPSTPTTFAFPQAGALWDEALPSAPGCIAPCLLDPPMKAV 180  
DB 121 SSPDEVLSTSMYFKQSPSTPTTFAFPQAGALWDEALPSAPGCIAPCLLDPPMKAV 180  
QY 181 PTVAGARFPLFHKPSPHPHPSPAGGHHGLGYDPTAAALSLPLGAAAAGSQAALRS 240  
DB 181 PTVAGARFPLFHKPSPHPHPSPAGGHHGLGYDPTAAALSLPLGAAAAGSQAALRS 240

QY 241 HPYGLPLAKRAAPLAPPPPLGLTPSPATASSLLGESPSLPSPSPSSSSSGGTCAVCGDNAA 300  
DB 241 HPYGLPLAKRAAPLAPPPPLGLTPSPATASSLLGESPSLPSPSPSSSSSGGTCAVCGDNAA 300  
QY 301 COHYGVRTCEGCKGFFKRTVQKNKXVCLANKNCVPDKRRNRRCQYCRFQKCLSVGMVKE 360  
DB 301 COHYGVRTCEGCKGFFKRTVQKNKXVCLANKNCVPDKRRNRRCQYCRFQKCLSVGMVKE 360  
QY 361 VVRTDSLKGRGRGLSPKPKSPLOQBPSPQSPSPPICMNVALVRLTDTSTPRDLDSRYC 420  
DB 361 VVRTDSLKGRGRGLSPKPKSPLOQBPSPQSPSPPICMNVALVRLTDTSTPRDLDSRYC 420  
QY 421 PTDOAAAGTDAEHVQFYNNLTASIDVRSWAEPKIPGFTDLPKEDDTLLIESAFLEFVL 480  
DB 421 PTDOAAAGTDAEHVQFYNNLTASIDVRSWAEPKIPGFTDLPKEDDTLLIESAFLEFVL 480  
QY 481 RLISIRNTAEDKVFVFCNGVLVLRLOCLRGFGWLDLSIKDFSLNQLSLNDIOALACLSAL 540  
DB 481 RLISIRNTAEDKVFVFCNGVLVLRLOCLRGFGWLDLSIKDFSLNQLSLNDIOALACLSAL 540  
QY 541 SMITERHGLKPKRVEELCNKITSSLKDHQSKGQALEPTESKVLGALVELRKICTLGLQR 600  
DB 541 SMITERHGLKPKRVEELCNKITSSLKDHQSKGQALEPTESKVLGALVELRKICTLGLQR 600  
QY 601 IFYKLKLEDLVSPSPSIIDKLFDLTLPF 626  
DB 601 IFYKLKLEDLVSPSPSIIDKLFDLTLPF 626

RESULT 6

AAW16398  
ID AAW16398 standard; protein; 626 AA.  
XX AAW16398;  
XX AAW16398;  
DT 08-SEP-1997 (first entry)  
XX Human neuron-derived orphan receptor NOR-1 protein.  
DE Human; neuron-derived orphan receptor; open reading frame; homology; rat;  
KW NOR-1; DNA binding domain; brain; neuronal; Alzheimer's disease;  
KW Parkinson's disease.  
XX Homo sapiens.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Domain 291..377  
FT /note= "DNA binding domain"  
XX JF09084585-A.  
XX 31-MAR-1997.  
XX 21-SEP-1995; 95JP-00242909.  
XX 21-SEP-1995; 95JP-00242909.  
XX (TERU ) TERUMO CORP.  
XX WPI; 1997-253003/23.  
XX N-PSDB; AAT73334.

XX Human neuron-derived orphan nuclear receptor gene - useful for diagnosing  
PT brain diseases such as Alzheimer's or Parkinson's disease or  
PT neuroblastoma.  
XX Claim 1; Page 5-8; 8pp; Japanese.  
XX This is the amino acid sequence of a novel human neuron-derived orphan  
CC nuclear receptor. The protein has high homology to the rat NOR-1  
CC receptor. The protein contains several DNA binding domains including P, D  
CC and A boxes which also have high homology to the rat NOR-1 DNA binding





PPA	(BAYU ) BAYLOR COLLEGE MEDICINE.
XX	(UNIW ) UNIV WASHINGTON.
XX	Mullican SE, Conneely OM, Milbrandt J;
XX	WPI; 2003-854017/79.
DR	GENBANK; CAA09763.
XX	Inhibiting proliferation of a hematopoietic cell, useful for treating or
PPT	preventing leukemia, comprises modulating the level or activity of nor-1
PT	and/or nur77 nuclear receptor.
XX	
PS	Disclosure; SEQ ID NO 17; 101pp; English.
XX	The present sequence is that of pig nor-1 nuclear receptor polypeptide
CC	CAA09763. Nor-1 has been identified as a molecular target for therapeutic
CC	intervention in the treatment of myeloid leukaemia. The invention
CC	provides methods of inhibiting the proliferation of a haematopoietic stem
CC	cell or a haematopoietic myeloid cell by modulating the level of a nor-1
CC	and/or nur77 nuclear receptor. This involves increasing the level of the
CC	receptor polypeptide or polynucleotide, e.g. by administration of a
CC	vector comprising the polynucleotide. A claimed method of treating
CC	leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a
CC	haematopoietic stem cell or myeloid cell. Also claimed are methods of
CC	identifying an upregulator of expression of nor-1 and/or nur77, of screening
CC	identifying a compound for the treatment of leukaemia, and of screening
CC	for a compound for treatment of leukaemia, and a mouse model for
CC	leukaemia comprising a mouse having defective nor-1 and nur77 nucleic
CC	acid sequences.
XX	
SQ	Sequence 643 AA;
	Query Match 93.9%; Score 3133.5; DB 7; Length 643;
	Best Local Similarity 92.4%; Pred. No. 2.le-208;
	Matches 594; Conservative 10; Mismatches 22; Indels 17; Gaps 3;
Qy	1 MPCVQAQYSPSPGSSYAAQYYSSEYTEIMNPDYKLTMDLGSTETATATSLPST 60
Db	1 MPCVQAQYSPSPGSSYAAQYYSSEYTEIMNPDYKLTMDLGSTETATATSLPST 60
Qy	61 FVEGYSSNVKLKPCVCYQM-Q---RPLIKVEEGRAPSYHHHHHHHHHHHQQ----- 110
Db	61 FMEGYSSNVKLKPCCLYQNPQSGRPLIKWEEGRAHYHHHHHHHHHHHQ----- 120
Qy	111 --CHOQPSTIPASSPEVLSTSMYFKQSPSTPTTPAPPQAGALWDEALPSAPGCIA 168
Db	121 PPOQQQSPISPPSGPEDEVLPSTSMYFKQSPSTPTTPVPPOQAGALWEDALPSAQGCIA 180
Qy	169 PGFLDPPMKAVPTVAGARFPLFHFKSPHPHPAPSAGGHILGYDPTAAAALSPLG-- 226
Db	181 PGFLDPPMKAVPTVAGARFPLFHFKSPHPHPAPSAGGHILGYDPTAAAALGLPGNA 240
Qy	227 ---AAAAAGSQAALESHPYLGLAKRAAPLAFFPLGLTSPSTASSLLGESPSLPSPSR 283
Db	241 AAAAAAAGSQAALEGHPYGLPLAKRAAALAFSPGLGLTTSTPTSSLGESPSLPSPPNR 300
Qy	284 SSSSGEGTCVACGDNAACOHTVTRTCECGCKFFKRTVKNAKYVCLANKNCVDKKERNR 343
Db	301 STASGEGTCVACGDNAACOHTVTRTCECGCKFFKRTVKNAKYVCLANKNCVDKKERNR 360
Qy	344 COYCREQKCLSVGMWKVEVRVTDLSKGRGRPLPSKPSPLOQEPSQSPSPSPPTCMNALV 403
Db	361 COYCRICQLSVGMWKVEVRVTDLSKGRGRPLPSKPSPLOQEPSQSPSPSPPVCMNALV 420
Qy	404 RALTDTSPRLDYSCYPTDDAAAGTDAEHVOQFNLLITASIDVRSRWAEKI PGFTDLPK 463
Db	421 RALTDTSPRLDYSCYPADQAAAGTDAEHVQVFNNLLITASIDVRSRWAEKIFGFTDLPK 480
Qy	464 EDOTLLIESAFLELVLRLSIRSNATAEDKFVFCNGLVLRLOCLRGEWLSDIKDFSLN 523
Db	481 EDOTLLIESAFLELVLRLSIRSNATAEDKFVFCNGLVLRLOCLRGEWLSDIKDFSRL 540
Qy	524 LQSLNLDIQALACLALSALMITSRHGLKEPKRVVELCNKITSSLDHKQSKQALEPTESKV 593









Db 61 FMEGYSPCEKLPSCLYQMPGSRPLIKMEEGREHYHHHHHHHHHHQ--QOPS 118  
Qy 117 IPPASSPEDEVLPTSTMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 176  
Db 119 IPPSPGPEDEVLPTSTMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 178  
Qy 177 MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 235  
Db 179 MKAVPMAAAARFPFIF-FKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 237  
Qy 236 AALESHPYGLPLAKRAAPLAPPLGLTPTASSLLGESPSLPSPSSSSSGSGTCAVC 295  
Db 238 AALEGHYPGLPLAKRTATLTTPPLGLTASPTASSLLGESPSLPSPSSSSSGSGTCAVC 297  
Qy 296 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDKRRNRCCYCRFQKCLSV 355  
Db 298 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDKRRNRCCYCRFQKCLSV 357  
Qy 356 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPICMNNALVRALTDSTPRDL 415  
Db 358 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPICMNNALVRALTDSTPRDL 417  
Qy 416 YSRICYPTDQAAAGTDAEHVQVQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAPL 475  
Db 418 YSRICYPTDQATAGTDAEHVQVQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAPL 477  
Qy 476 ELFVLRLSIRSNATDAEDKVFVFCNGLVHLRLQCLRGFGWLDISKDFSLNLSLNDIQALA 535  
Db 478 ELFVLRLSIRSNATDAEDKVFVFCNGLVHLRLQCLRGFGWLDISKDFSLNLSLNDIQALA 537  
Qy 536 CLSALSMITERHGLKEPKVEELCNKITSSLDKHOSKGOALEPTESKVLGALVELRKICT 595  
Db 538 CLSALSMITERHGLKEPKVEELCNKITSSLDKHOSKGOALEPTESKVLGALVELRKICT 597  
Qy 596 LGLQRIFFYLKLEDLVSPSPSIDKFLDITLPP 626  
Db 598 LGLQRIFFYLKLEDLVSPSPSIDKFLDITLPP 628

RESULT 15  
ID ADF28880  
AC ADF28880 standard; protein; 628 AA.  
XX ADF28880;  
XX  
XX 12-FEB-2004 (first entry)  
XX  
XX Rat nor-1 nuclear receptor polypeptide JC2493.  
XX  
XX Rat; nor-1; nuclear receptor; receptor; leukaemia; gene therapy;  
XX Cystostatic; haematopoietic cell.  
XX Rattus sp.  
XX OS  
XX WO2003088812-A2.  
XX 30-OCT-2003.  
XX 15-APR-2003; 2003WO-US011804.  
XX 17-APR-2002; 2002US-0373238P.  
XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX (UNIW ) UNIV WASHINGTON.  
XX Mullican SE, Conneely OM, Milbrandt J;  
XX WPI; 2003-854017/79.  
XX Inhibiting proliferation of a hematopoietic cell, useful for treating or  
XX preventing leukemia, comprises modulating the level or activity of nor-1  
XX and/or nur77 nuclear receptor.

XX  
PS  
XX  
CC The present sequence is that of rat nor-1 nuclear receptor polypeptide  
CC JC2493. Nor-1 has been identified as a molecular target for therapeutic  
CC intervention in the treatment of myeloid leukaemia. The invention  
CC provides methods of inhibiting the proliferation of a haematopoietic stem  
CC cell or a haematopoietic myeloid cell by modulating the level of the  
CC and/or nur77 nuclear receptor. This involves increasing the level of a  
CC receptor polypeptide or polynucleotide, e.g. by administration of a  
CC vector comprising the polynucleotide. A claimed method of treating  
CC leukaemia comprises modulating a nor-1 and/or nur77 nuclear receptor in a  
CC haematopoietic stem cell or myeloid cell. Also claimed are methods of  
CC identifying an upregulator of expression of nor-1 and/or nur77, of  
CC identifying a compound for the treatment of leukaemia, and of screening  
CC for a compound for treatment of leukaemia, and a mouse model for  
CC leukaemia comprising a mouse having defective nor-1 and nur77 nucleic  
CC acid sequences.  
XX  
SQ Sequence 628 AA;  
Query Match 92.1%; Score 3074; DB 7; Length 628;  
Best Local Similarity 92.2%; Pred. No. 2.7e-204;  
Matches 582; Conservative 12; Mismatches 29; Indels 8; Gaps 4;  
Qy 1 MPCVQAQYSPSPGSSVAAOTYSSEYTTIMNDYTKLTMDLSTETATATSLPSIST 60  
Db 1 MPCVQAQYSPSPGSSVAAOTYSSEYTTIMNDYTKLTMDLSTETATATSLPSIST 60  
Qy 61 FVEGYSNVELKPSVCVQM---ORPLIKVEEGRAPSYHHHHHHHHHHHHHHHHHHHH 116  
Db 61 FMEGYSPCEKLPSCLYQMPGSRPLIKMEEGREHYHHHHHHHHHHHHHHHHHHHH 118  
Qy 117 IPPASSPEDEVLPTSTMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 176  
Db 119 IPPSPGPEDEVLPTSTMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPP 178  
Qy 177 MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 235  
Db 179 MKAVPMAAAARFPFIF-FKPSPPHPPAPSPAGGHHGLGYDPTAAALSLPLGAAAAGSOA 237  
Qy 236 AALESHPYGLPLAKRAAPLAPPLGLTPTASSLLGESPSLPSPSSSSSGSGTCAVC 295  
Db 238 AALEGHYPGLPLAKRTATLTTPPLGLTASPTASSLLGESPSLPSPSSSSSGSGTCAVC 297  
Qy 296 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDKRRNRCCYCRFQKCLSV 355  
Db 298 GDNAAQOHYGVRTCEGCKGFFKRTVQKNAKYVCLANKNCVDKRRNRCCYCRFQKCLSV 357  
Qy 356 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPICMNNALVRALTDSTPRDL 415  
Db 358 GMVKEVVRTDSLKGRGRPLSPKPSPLQOEPSQSPSPSPICMNNALVRALTDSTPRDL 417  
Qy 416 YSRICYPTDQAAAGTDAEHVQVQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAPL 475  
Db 418 YSRICYPTDQATAGTDAEHVQVQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIESAPL 477  
Qy 476 ELFVLRLSIRSNATDAEDKVFVFCNGLVHLRLQCLRGFGWLDISKDFSLNLSLNDIQALA 535  
Db 478 ELFVLRLSIRSNATDAEDKVFVFCNGLVHLRLQCLRGFGWLDISKDFSLNLSLNDIQALA 537  
Qy 536 CLSALSMITERHGLKEPKVEELCNKITSSLDKHOSKGOALEPTESKVLGALVELRKICT 595  
Db 538 CLSALSMITERHGLKEPKVEELCNKITSSLDKHOSKGOALEPTESKVLGALVELRKICT 597  
Qy 596 LGLQRIFFYLKLEDLVSPSPSIDKFLDITLPP 626  
Db 598 LGLQRIFFYLKLEDLVSPSPSIDKFLDITLPP 628

Search completed: March 12, 2005, 17:48:10  
Job time : 173 secs

1	3337	100.0	697	4	US-09-949-016-9660	Sequence 9650, Ap
2	3316	99.4	626	4	US-09-949-016-6776	Sequence 6776, Ap
3	1751	52.5	598	4	US-09-949-016-6699	Sequence 6699, Ap
4	1751	52.5	617	4	US-09-949-016-10338	Sequence 10338, A
5	1367	41.0	641	4	US-09-949-016-10263	Sequence 10263, A
6	659.5	19.8	300	3	US-09-277-078-2	Sequence 2, Appli
7	533.5	16.0	533	1	US-07-952-800-2	Sequence 2, Appli
8	530.5	15.9	448	4	US-09-949-016-8178	Sequence 8178, Ap
9	530.5	15.9	462	2	US-08-592-383-2	Sequence 2, Appli
10	530.5	15.9	462	2	US-08-095-7288-4	Sequence 4, Appli
11	530.5	15.9	462	5	PCT-US92-02320A-4	Sequence 4, Appli
12	530.5	15.9	525	3	US-08-764-870-7	Sequence 7, Appli
13	530.5	15.9	525	3	US-08-980-115-7	Sequence 7, Appli
14	530.5	15.9	533	4	US-08-216-592A-4	Sequence 4, Appli
15	530.5	15.9	577	4	US-09-949-016-11572	Sequence 11572, A
16	527.5	15.8	462	6	5171671-2	Sequence 11572, A
17	527.5	15.8	462	6	5171671-2	Sequence 11572, A
18	525.5	15.7	446	1	US-07-952-800-4	Sequence 4, Appli
19	524.5	15.7	448	4	US-08-216-592A-2	Sequence 2, Appli
20	516	15.5	403	2	US-08-592-383-4	Sequence 4, Appli
21	502	15.0	454	4	US-09-949-016-11665	Sequence 11665, A
22	502	15.0	462	1	US-08-336-408B-2	Sequence 2, Appli
23	502	15.0	462	3	US-08-764-870-6	Sequence 6, Appli
24	502	15.0	462	3	US-08-980-115-6	Sequence 6, Appli
25	502	15.0	462	4	US-10-329-668-10	Sequence 10, Appli
26	502	15.0	462	5	PCT-US91-00399-2	Sequence 2, Appli
27	500.5	15.0	503	4	US-09-949-016-9851	Sequence 9851, Ap

```
Db 372 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKKCPVDKRRNRQYCRFOKCLSVGMVKE 431
Qy 361 VVRTDSLKGRGRPLSPKPSKPLQOEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Db 432 VVRTDSLKGRGRPLSPKPSKPLQOEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 491
Qy 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 492 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 551
Qy 481 RLSIRSTAEDKVFVFCNGLVLRHQLQCLRGFGEWLDSIKDPSLNQSLNLDIOALACLSAL 540
Db 552 RLSIRSTAEDKVFVFCNGLVLRHQLQCLRGFGEWLDSIKDPSLNQSLNLDIOALACLSAL 611
Qy 541 SMITERHGLKEPKRVEELCNKITSSLDKHSKQGALEPTESKVLGALVELRKICTLGLQR 600
Db 612 SMITERHGLKEPKRVEELCNKITSSLDKHSKQGALEPTESKVLGALVELRKICTLGLQR 671
Qy 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
Db 672 IFYKLEDLVSPPSIIDKFLDTLPF 697
```

## RESULT 2

```
US-09-949-016-6776
; Sequence 6776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6776
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6776
```

```
Query Match 99.4%; Score 3316; DB 4; Length 626;
Best Local Similarity 99.4%; Pred. No. 2.4e-247;
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTEITATATSLPSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTEITATATSLPSIST 60
Qy 61 FVEGYSNNYELKPSQVQWRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSNNYELKPSQVQWRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Qy 121 SSPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
Db 121 SSPDEVLSTSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
Qy 181 PTVAGARFPLFHFKEPSPHPHPPSPAGGHLGVDPTAAALSLPLGAAAAGSQAALLES 240
Db 181 PTVAGARFPLFHFKEPSPHPHPPSPAGGHLGVDPTAAALSLPLGAAAAGSQAALLES 240
Qy 241 HPYGLPLAKRAAPLAFPLGLTSPASSLLGESPSLPGPPSSSSSGEGTCAVCGDNA 300
Db 241 HPYGLPLAKRAAPLAFPLGLTSPASSLLGESPSLPGPPSSSSSGEGTCAVCGDNA 300
```

```
Qy 301 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKKCPVDKRRNRQYCRFOKCLSVGMVKE 360
Db 301 COHGVRTCEGCKGFFKRTVQKNAKYVCLANKKCPVDKRRNRQYCRFOKCLSVGMVKE 360
Qy 361 VVRTDSLKGRGRPLSPKPSKPLQOEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Db 361 VVRTDSLKGRGRPLSPKPSKPLQOEPSQSPSPPICMNNALVRALTDSTPRDLDSRYC 420
Qy 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 421 PTQAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Qy 481 RLSIRSTAEDKVFVFCNGLVLRHQLQCLRGFGEWLDSIKDPSLNQSLNLDIOALACLSAL 540
Db 481 RLSIRSTAEDKVFVFCNGLVLRHQLQCLRGFGEWLDSIKDPSLNQSLNLDIOALACLSAL 540
Qy 541 SMITERHGLKEPKRVEELCNKITSSLDKHSKQGALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLDKHSKQGALEPTESKVLGALVELRKICTLGLQR 600
Qy 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
Db 601 IFYKLEDLVSPPSIIDKFLDTLPF 626
RESULT 3
US-09-949-016-6699
; Sequence 6699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6699
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6699
Query Match 52.5%; Score 1751; DB 4; Length 598;
Best Local Similarity 54.5%; Pred. No. 1.1e-126;
Matches 363; Conservative 74; Mismatches 121; Indels 108; Gaps 18;
Qy 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTEITATATSLP 56
Db 1 MPCVQAQYSPSPGSSYAAQYTSSEYTTIMNPDYTKLTMDLGSTEITATATSLP 58
Qy 57 SISTFVEGYSNNYELKPSQVQWRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHH 112
Db 59 SPSTFMDNVSTGYDVKPPCLYQMLPSGQSSIKVEDIQMHNYQQHSH----- 105
Qy 113 QQPSTPPASSDEVLPLP-STSMYFKQSPSTPTTAPFPQAGALWDEALPSAPGCIAPGP 171
Db 106 -----LPPQS-----EEMWPHSGSVYIKPSPPTTTPGQVQHSIPMWD----- 148
Qy 172 LLDPPMKAVPT-----VAGARFPLFHFKEPSPHPH-----PSAPG 207
Db 149 LHNHFQNYVATHTMIEQRKTPVSRSLSLFSFKQSPGPTPVSSQCMRFDGDLHPMNP 208
Qy 208 GHHLGYDPTAAALSLPLGAAAAGSQAALLESHPYGLP-LAKRAAPLAFPLGLTSPPT 266
Db 209 SHHV-----VDQGTFAVFNPIRKPSMGFPGLQI-----CH 239
```



Db 153 GCYFGLSGVDALSSSGSYGSCAPSPSTFQPPQLSPWDGSGFHGFSQTYEG 212  
QY 154 -ALMDEALPSAPGCIAPGLDPPMKAVPTVAGARFPLFHFKPSPPPAPSPAGGHLG 212  
Db 213 LRANTEQLPKASG-----PPQ-----PPAFFSPPTGPSES----- 244  
QY 213 YDPTAAALSL-PLGAAAGSQAALSHYPGLPLAKRAAPLAPPLPLGLTPSPASSLL 271  
Db 245 ---LAQSPKLPFSPQATHQLG-----EGESYSM-----PTAPPLG-----APTSPHLE 284  
QY 272 G-----ESPGLPPSPSSSSG--EGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNK 325  
Db 285 GSGILDTP-VTSYKARSGAPGSEGRCAVCGDNASCQHYGVRTCEGCKGFFKRTVQKNK 343  
QY 326 YVCLANKCPVDKRRNRNCOYCRFOKCLSVGMVKEVVRTDSLKGRGLPSKPKSPLOQE 385  
Db 344 YICLANKDCPVDKRRNRNCOYCRFOKCLAVGMVKEVVRTDSLKGRGLPSKPKQP----- 399  
QY 386 PSQSPSPSPICMNNALVREALTDSTP--RDLDSYVCPTDQAAAG-TDAEHVQQFYNLIT 442  
Db 400 -----PDASPANLLTSLVRAHLDGSPSTAKLDYSKFQELVLPHFGKEDAGDVQOQFYDLS 454  
QY 443 ASIDVSRSWAEPKIPGTDLPKEDQTLIESAFLELVFLRLSTRSNTAEDKFVFCNGLVLH 502  
Db 455 GSLEVRKWAEPKIPGFAELSPADQDLLESFALELVFLRLAVRSKPGCKLIFCSGLVLH 514  
QY 503 RLQCLRGFEWLDISKDFNSLQSLMDLQALACLALSMTIRHCLKEPKEVEELCNKI 562  
Db 515 RLQCARGFWDISILAFSRLSHSLVLLVDPAFACLSALVLIITDRHGLQEPKRVESLQNKI 574  
QY 563 TSSLKDHQSGQALEPTE----SKVLGALVELRKICTLGLQIFVYKLEDLVSPSSIIDK 618  
Db 575 ASCLKEHVA-AVAGEPQASCLSRLLGKLPBLTCTQGLQIFVYKLEDLVSPSSIIDK 633  
QY 619 LFLDTLPF 626  
Db 634 IFMDTLPF 641

RESULT 6  
US-09-277-078-2  
; Sequence 2, Application US/09277078  
; Patent No. 6312949  
; GENERAL INFORMATION:  
; APPLICANT: Sakurada, Kazuhiro  
; APPLICANT: Palmer, Theo  
; TITLE OF INVENTION: REGULATION OF TYROSINE HYDROXYLASE  
; TITLE OF INVENTION: EXPRESSION  
; FILE REFERENCE: 07251/031001  
; CURRENT APPLICATION NUMBER: US/09/277,078  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-277-078-2

Query Match 19.8%; Score 659.5; DB 3; Length 300;  
Best Local Similarity 41.6%; Pred. No. 6.2e-43;  
Matches 151; Conservative 41; Mismatches 74; Indels 97; Gaps 13;  
QY 1 MPCVQAQYSSPPGSSVAAQTY-----SSEYTTETMNDYTKLMDLSTGSTITATATSLP 56  
Db 1 MPCVQAQYSSPPGSSVAAQTY-----SSEYTTETMNDYTKLMDLSTGSTITATATSLP 58  
QY 57 SISTFVEGYSNVELKPSVCYQM-----QRPLIKVEGRAPSYHHHHHHHHHHHHHQQOH 112  
Db 59 SFSTFMDNSTGYDVKPCLYQMLSGQSSIKVEDIQMNYQQHSH----- 105

QY 113 QQPSIIPASSPEDEVLP-STSMYFKOSPSTTPTTAPFPQAGALMDEALPSAPGCIAPGP 171  
Db 106 ----LPPQS---BEMPHSGSVYTKFSSPTPTPGFQVQHSFPMWDD-----PGS 148  
QY 172 LLDPKPKAVPT-----VAGARFPLFHFKPSPPPHPPA-----PSPAG 207  
Db 149 LNHFNQNYVATTMI EOKTVPVSRLLSLFSFKQSPGPTFVSQCMRFDGFLHVPNMPPEAG 208  
QY 208 GHILGYDPTAAALSLPLGAAAGSQAALSHYPGLP-LAKRAAPLAPPLPLGLTPSPT 266  
Db 209 SHHV-----LDQTFAPVNPPIRKPSMGFFGLQI-----GH 239  
QY 267 ASSILGSSPLSPPPSSSSSGEGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNK 326  
Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACQHYGVRTCEGCKGFFKRTVQKNK 297  
QY 327 VCL 329  
Db 298 VCL 300

RESULT 7  
US-07-952-800-2  
; Sequence 2, Application US/07952800  
; Patent No. 5403925  
; GENERAL INFORMATION:  
; APPLICANT: OZATO, KEIKO  
; TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE  
; TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND  
; STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000  
; CITY: SAN FRANCISCO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/952,800  
; FILING DATE: 19920928  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-21-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 533 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-952-800-2

Query Match 16.0%; Score 533.5; DB 1; Length 533;  
Best Local Similarity 29.7%; Pred. No. 6.8e-33;  
Matches 151; Conservative 70; Mismatches 189; Indels 99; Gaps 17;  
QY 118 PRASSPEDEVLPSTSMYFKOSPSTTPTTAPFPQAGALMDEALPSAPGCIAPGLDPPM 177  
Db 83 PDSSSNP--LP-----QGVPPSPPPGPPPLPSTA-----PSLGGSGGAPP-----PPM 124  
QY 178 KAVPTVAGARFPLFHFKPSPPPHPPAPSPAGGHLGYDPTAAALSLPLGAAAGSQAAL 237  
Db 125 P--PPPLGSPFPVVISSSMGSPGLPPAPPFGFSGPVSSPQINSTVSLPGSGG----- 174

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPSPPTASSLLGESPSLPSPPSSSSSGEGTCVAGCD 297  
Db 175 -----PPEDVFP-----PVLGVRGLHCPPPGPGGAGKRLCAICGD 210  
QY 298 NAACQHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCPVDKRRNRQCYCRFQKCLSVGM 357  
Db 211 RSSGKHGYVSGEGCKGFFKRTKDLTVYSCRDNDKCTVDKQRNRQCYCRQKCLATGM 270  
QY 358 VKEVVRTSLKRRGRLSPKSPLOQBPSPQSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 406  
Db 271 KREAVOEERQSG-----KQDGG-DGEGAGGAPPEMPVDRIEAEAEVQKSDQGVGP 322  
QY 407 -----TDSPTPRLDYSRYCPTDQAAAGTDAEHVQOQFYNNLTASIDVSRSAEKIIFGFTDL 461  
Db 323 GGTGSGSGSPND-----PVTNICOAD-----KQLFTLV-----EWAKRIHFSSSL 363  
QY 462 PKEDQTLTIESAFLEFVLRLSIRNSNTAEDKEFVFCNGLVHLRLQC-LRGFGWMLDSI-KD 519  
Db 364 PLDDQVILLRAGWNELLTASFHSRSDIVDRDGILLATGLHVRHNSAHSAGVGAIFDRVLTE 423  
QY 520 FSLNQLSLNLDIQALACLALSLSMIT-ERHGLKEPKRVEELCNKITSSLSKDH-QSKGOALE 577  
Db 424 LVSKNRDMRMKTELGLCLRAITLFPNDAGLSNPSEVEVLEKRVYASLETYCKQKYPBQQ 483  
QY 578 PTESKVLGALVELRICKTGLQRIFFVLKL 606  
Db 484 GRFALLRLRLPALRSIGLKCLEHLFFFKL 512

## RESULT 8

US-09-949-016-8178  
; Sequence 8178, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8178  
; LENGTH: 448  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(448)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-949-016-8178

Query Match 15.9%; Score 530.5; DB 4; Length 448;  
Best Local Similarity 34.5%; Pred. No. 9.3e-33;  
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;

QY 203 PSPAGGHGLGYDPTAAALSLP--IGAAAAGSQAALLESHPYGLPLAKRAAPLAPPLG 260  
Db 14 PTPGGHLNGY-PVPYPAFFFPPLGLGSLPPG-----ALTTLQHLFVSGYS----- 59  
QY 261 LTPSP-----TASSLLGE-SPSPSPSSSSSGEGTCVAGCNAACQHYGVRTCEGCKGF 315  
Db 60 -TPSPATTETQSSSEIEIVSPSPSPPLPRI---YKPCFVCQDKSGYHYGVSAECGCKGF 116  
QY 316 FKRTVQNAKYVCLANKNCPVDKRRNRQCYCRFQKCLSVGMVKEVVRTSLKRRGRPL 375  
Db 117 FRRSIQKNVYCHRDKNKCIINKVTNRQCYCRQKCEVGVGSKESVRNDRNK----- 169

QY 376 SKPKSPLOQBPSP 432  
Db 170 -KKKEVPPECSSESYTLTPEVGELIEKVRKAHQETFPALCQLGKTYTNNSSQORVSLDID 228  
QY 433 HVQOQFYNNLTASIDVSRSAEKIIFGFTDLPKEDQTLTIESAFLEFVLRLSIRNSNTAEDK 492  
Db 229 LWDKESELSKCIITKTEFAKQLPGFTTLTIADQITLLKAAACLDILILICITRYTPEQDT 288  
QY 493 FVFCNGLVHLRLQC-LRGFGWMLDSI-KD 519  
Db 289 MTFSDGLTLNRQTMHNAFGPLTDLVFAFANQLLPLEMDDAETGLLSAICLCGRQDLE 348  
QY 551 EPRKVEELCNKITSSLSKDHQSKGOALEP-TESTKVLGALVELRICKTGLQRIFFVLKL 609  
Db 349 QPDRVDMLEQLEALKVYVRKRPSRPHMFMMLMKITDLRSISAKGAERVITLKWEP 408  
QY 610 VSPPSIIDKLF-----LDTL 624  
Db 409 GSMPLIQEMLENSGLDTL 428

## RESULT 9

US-08-592-383-2  
; Sequence 2, Application US/08592383  
; Patent No. 5830760  
; GENERAL INFORMATION:  
; APPLICANT: Teal, S. and S.J. Collins  
; TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Retinoic Acid Recep  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness  
; STREET: 2800 Pacific First Centre, 1420 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage  
; COMPUTER: IBM PC/386 Compatible  
; OPERATING SYSTEM: MS-DOS 4.01  
; SOFTWARE: Word for Windows 5.01-t  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/592,383  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/099,242  
; FILING DATE: July 28, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brogerick, Thomas F.  
; REGISTRATION NUMBER: 31,332  
; REFERENCE/DOCKET NUMBER: FHCR-1-7190  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 1-206-682-8100; 1-206-224-0709(direct)  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: polypeptide  
; DESCRIPTION: sequence encoded by SEQ. ID. NO. 1  
US-08-592-383-2

Query Match 15.9%; Score 530.5; DB 2; Length 462;  
Best Local Similarity 34.5%; Pred. No. 9.6e-33;  
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;  
QY 203 PSPAGGHGLGYDPTAAALSLP--IGAAAAGSQAALLESHPYGLPLAKRAAPLAPPLG 260  
Db 9 PTPGGHLNGY-PVPYPAFFFPPLGLGSLPPG-----ALTTLQHLFVSGYS----- 54



261 LTPSP-----TASSLLGE-SPSLPSPSSSSGEGTCAVCGDNAACOHYGVRTCEGCKGF 315  
 55 -TPSPATITQSSSEIIVPSPSPPLPRI--YKPCFVQODKSSGYHYGVSACEGCKGF 111  
 316 FKRTVQNAKYVCLANKCPVDKRRNRNCOYCRFOKCLSVGMVKEVVVTRTSLKGRGRPL 375  
 112 FRSIQKNWYVTHRDKNKIINKVTRNCOYCRLOKCFEVMGSKESVRNDRNK-----164  
 376 SKPSPLOQEPSQSPSPSPPICMNNALVR-ALTDSTPRDLDSYRCP--TDQAAAGTDAE 432  
 165 -KKKEVPKPCSESYTLTPEVGELIEKVRKAHOETFPALCQKGYTYTNNSEQRVSLDID 223  
 433 HVQOYNLLTASIDVSRWAELKIPGFTDLPKEDQTLIESAFLEFVLRLSIRNTAEDK 492  
 224 LWDKFSLSKCIKIVTVEFAKQLPFGFTTLTADQITLLKAACLDLILIRICTRYTPEQDT 283  
 493 FVFCNGVLVHRLQCLR-GFGEWLDISKDFSLNLSQNLDIQALACLSALSMIT-ERHGLK 550  
 284 MTFSDGLTLNRTQMHNAGFGLTDLVFAFANQLLPLEMDDAETGLLSAICLICGDRQDLE 343  
 551 EPRKVEELCNKITSSLDKSHQSGQALEP-TESKVLGALVELRKICTLGLQRIFYLKLEDL 609  
 344 QPDRVDMLEPQLEALKVYVRKRRPRSPHMPKMLMKITLRSISAKGAERVITLKMEIP 403  
 610 VSPPSIIDKLF-----LDTL 624  
 404 GSNPPLIQEMLENSGLDITL 423

RESULT 10  
 US-08-095-728B-4  
 ; Sequence 4, Application US/08095728B  
 ; Patent No. 583642  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DMITROVSKY, ETHAN  
 ; APPLICANT: WARRELL JR, RAYMOND P  
 ; APPLICANT: MILLER JR, WILSON H  
 ; APPLICANT: FRANKEL, STANLEY  
 ; TITLE OF INVENTION: METHODS FOR THE DETECTION AND  
 ; TITLE OF INVENTION: TREATMENT OF ACUTE PROMYELOCYTIC LEUKEMIA (APL)  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: COOPER & DUNHAM LLP  
 ; STREET: 1185 AVENUE OF THE AMERICAS  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.24  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/095,728B  
 ; FILING DATE: 21-JUL-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/673,838  
 ; FILING DATE: 22-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: WHITE, JOHN P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 38694-A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0525  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 462 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-08-095-728B-4  
 Query Match 15.9%; Score 530.5; DB 2; Length 462;  
 Best Local Similarity 34.5%; Pred. No. 9.6e-33;  
 Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;  
 QY 203 PSPAGGHHGYDPTAAALSLP--LGAASAAASQAALSHHPYGLPLAKRAAPLAPPLG 260  
 DB 9 PTGGGHLNGY-PVPPYAPFPFPLGGLSPG-----ALTTLOHQLFVSGYS-----54  
 QY 261 LTPSP-----TASSLLGE-SPSLPSPSSSSGEGTCAVCGDNAACOHYGVRTCEGCKGF 315  
 DB 55 -TPSPATITQSSSEIIVPSPSPPLPRI--YKPCFVQODKSSGYHYGVSACEGCKGF 111  
 QY 316 FKRTVQNAKYVCLANKCPVDKRRNRNCOYCRFOKCLSVGMVKEVVVTRTSLKGRGRPL 375  
 DB 112 FRSIQKNWYVTHRDKNKIINKVTRNCOYCRLOKCFEVMGSKESVRNDRNK-----164  
 QY 376 SKPSPLOQEPSQSPSPSPPICMNNALVR-ALTDSTPRDLDSYRCP--TDQAAAGTDAE 432  
 DB 165 -KKKEVPKPCSESYTLTPEVGELIEKVRKAHOETFPALCQKGYTYTNNSEQRVSLDID 223  
 QY 433 HVQOYNLLTASIDVSRWAELKIPGFTDLPKEDQTLIESAFLEFVLRLSIRNTAEDK 492  
 DB 224 LWDKFSLSKCIKIVTVEFAKQLPFGFTTLTADQITLLKAACLDLILIRICTRYTPEQDT 283  
 QY 493 FVFCNGVLVHRLQCLR-GFGEWLDISKDFSLNLSQNLDIQALACLSALSMIT-ERHGLK 550  
 DB 284 MTFSDGLTLNRTQMHNAGFGLTDLVFAFANQLLPLEMDDAETGLLSAICLICGDRQDLE 343  
 QY 551 EPRKVEELCNKITSSLDKSHQSGQALEP-TESKVLGALVELRKICTLGLQRIFYLKLEDL 609  
 DB 344 QPDRVDMLEPQLEALKVYVRKRRPRSPHMPKMLMKITLRSISAKGAERVITLKMEIP 403  
 QY 610 VSPPSIIDKLF-----LDTL 624  
 DB 404 GSNPPLIQEMLENSGLDITL 423

RESULT 11  
 PCT-US92-02320A-4  
 ; Sequence 4, Application PC/TUS9202320A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sloan-Kettering Institute, For Cancer Research  
 ; TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: COOPER & DUNHAM  
 ; STREET: 30 ROCKEFELLER PLAZA  
 ; CITY: NEW YORK  
 ; STATE: NEW YORK  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10112  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.24  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US92/02320A  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 675,084  
 ; FILING DATE: 22-MAR-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 673,838  
 ; FILING DATE: 22-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: WHITE, JOHN P  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 38694-PCT  
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550  
TELEFAX: (212) 644-0525  
TELEX: (212) 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-02320A-4

Query Match 15.9%; Score 530.5; DB 5; Length 462;  
Best Local Similarity 34.5%; Pred. No. 9.6e-33;  
Matches 152; Conservative 64; Mismatches 181; Indels 43; Gaps 14;

QY 203 PSPAGCHLGYDPTAAALSLP--IGAAAAAGSQAALSHYPYGLPLAKRAAPLAPPPILG 260  
DB 9 PTPGGCHLNGY-PVPYPAFFPPMLGGLSPG---ALTTLQHLPVSGYS-----54

QY 261 LTPSP----TASSLLGE-SPSLPSPSRSSSSGEGTCAVCGDNACQAHYGVRTCEGCKGF 315  
DB 55 -TPSPATITETOSSEIEIVSPSPPPPLPRI--YKPCFVCQDKSGYHYGVSAACGCKGF 111

QY 316 FKRTQVNAKYVCLANKNCVDRNRNRCQYCRFOKCLSVGMVKEVVRTDSLKRRGRRLP 375  
DB 112 FRSIQKNVYVYCHRDKNKCIINKVTRNRCQYCRLOKCEVGVGSKESVRNRNK-----164

QY 376 SKPKSPLOQEPSQSPSPSPPICMNALVR-ALTDSTPRDLVSRYP--TDQAAAGTDAE 432  
DB 165 -KKEVPKPECSYSLTPEVCELIEKVKAKHETFPALCQLGKYVNTNNSRQVSLDID 223

QY 433 HVQYQYNLLTASIDVSRSWAEKIPTGDTLPKEDOTLLTESALELFLVRLSIRSNATDK 492  
DB 224 LWDKFSLSLTKCIITVTEFAKLPGLPTTLTIADQITLLKAAACLDILIRICTRYTPEQDT 283

QY 493 FVFCNGLVLRHQCLLR-GFGEMLDSIKDFSLMLQSLNLDIQALACLSALSMIT-ERHGLK 550  
DB 284 MTFSGDLTLNRTOMHNAFGPLTDLVFAFANQLLPLEWDAETGLLSAICLICGRQDL 343

QY 551 EPRVVEELCNKTTSLKHQSKQALEP--TESKVLGALVELRKICTLGLQRIIFYLKJEDL 609  
DB 344 QPDRVDMLOEPLLEALKVYVRRRPRPHMPFPMKMLKITDLRSISAKGAERVITLKMEIP 403

QY 610 VSPPSIIDKLP-----LDTL 624  
DB 404 GMPPELIQEMLENSGLDTL 423

RESULT 12

US-08-764-870-7  
Sequence 7, Application US/08764870  
Patent No. 6236946  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Thomas S  
APPLICANT: Baxter, John D  
APPLICANT: Fletterick, Robert J  
APPLICANT: Wagner, Richard L  
APPLICANT: Kushner, Peter J  
APPLICANT: Apriletti, James W  
APPLICANT: West, Brian  
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand  
NUMBER OF INVENTION: Binding Domains  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward  
STREET: Five Palo Alto Square, 3000 El Camino Real  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,870  
FILING DATE: 13-DEC-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,540  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,543  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,606  
FILING DATE: 14-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakamura, Jackie N  
REGISTRATION NUMBER: 35,966  
REFERENCE/DOCKET NUMBER: UCAL-246/01US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)843-5000  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 525 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-764-870-7

Query Match 15.9%; Score 530.5; DB 3; Length 525;  
Best Local Similarity 29.2%; Pred. No. 1.1e-32;  
Matches 150; Conservative 68; Mismatches 187; Indels 109; Gaps 17;

QY 118 PPASSPEDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGCLLDPPM 177  
DB 75 PDSSSPNP--LP-----QGVPPSPPGPLPPSTA-----PTLGGSGAAPP---PPM 116

QY 178 KAVPTVAGARPPLPHFKSPSPHPAPSPAGHHGLGYDPTAAALSLPLGAAAAAGSQA 237  
DB 117 P--PPPLGSPFPFVISSSMGSPGLPPAPPGFSGPVSSPQINSTVSLPGGGSG-----166

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTSTASSLLGESPLSPBSRSSSSGEGTCAVCGD 297  
DB 167 -----PPEDVKP-----PVLGVRGLHCPPPGPGGAGKRLCAICGD 202

QY 298 NAAQCHYGVRTCEGCKGFKRTVQKNKYVCLANKNCVDRNRNRCQYCRFOKCLSVGM 357  
DB 203 RSSGHYGVYSCGCKGFFKRTIRKDLTYSCEDNKDCTVDKQRNRCQYCRFOKCLATGM 262

QY 358 VKEVVRTDSLKGR-----RGRLPSPKSPLOQEPSQSPSPSPPICMNALVRALTD- 408  
DB 263 KREAVQEEQRQKDKDGEACAGAPEE-----MPVDRILEAEALAVEQKSDQ 309

QY 409 -----STPRDLDSRYCPTDQAAAGTDAEHVQOQFYNNLLTASIDVSRSWAEKIP 456  
DB 310 GVEGPGGTGGSGSSPND-----PVTNICQAAD-----KQLFTLV-----EWAKRIP 350

QY 457 GFTDLPKEDQTLLESAFLELFLVRLSIRSNATAEDKFCNGLVLRHQLC-LRGFGCEWLD 515  
DB 351 HFSSLPDLDVILLRAGWNELLIAFSHRSIDVRGILLATGLHVRNSAHAGVGAIFD 410

QY 516 SI-KDFSLNLOSINLDIQALACLSALSMIT-ERHGLKEPKRVVEELCNKTTSLKHQ- 572  
DB 411 RVLTELVSVMKDRMDKTELGLCLRAIILFNPDAKGLSNPSEVEVLEKRYASLETYCKQK 470

QY 573 GQALEPTESKVLGALVELRKICTLGLQRIIFYLKL 606  
DB 471 YPEOQGRFAKGLLRPLALRSIGLKCLEHLFFPKL 504

RESULT 13  
US-08-980-115-7

; Sequence 7, Application US/08980115

; Patent No. 6266622

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Thomas S.

; APPLICANT: Baxter, John D.

; APPLICANT: Fletcher, Robert J.

; APPLICANT: Wagner, Richard L.

; APPLICANT: Kushner, Peter J.

; APPLICANT: Apriletti, James W.

; APPLICANT: West, Brian L.

; APPLICANT: Shiao, Andrew K.

; TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS

; FILE REFERENCE: UCAL-246/02US

; CURRENT APPLICATION NUMBER: US/08/980,115

; CURRENT FILING DATE: 1997-11-26

; EARLIER APPLICATION NUMBER: 08/964,870

; EARLIER FILING DATE: 1996-12-13

; EARLIER APPLICATION NUMBER: 60/008,606

; EARLIER FILING DATE: 1995-12-14

; EARLIER APPLICATION NUMBER: 60/008,543

; EARLIER FILING DATE: 1995-12-13

; EARLIER APPLICATION NUMBER: 60/008,540

; EARLIER FILING DATE: 1995-12-13

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 525

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; LOCATION: (292)..(523)

; OTHER INFORMATION: minimal ligand binding domain

US-08-980-115-7

Query Match 15.9%; Score 530.5; DB 3; Length 525;

Best Local Similarity 29.2%; Pred. No. 1.1e-32;

Matches 150; Conservative 68; Mismatches 187; Indels 109; Gaps 17;

QY 118 PPASPEDEVLPSTSMYKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGGLDPPM 177

DB 75 PDSSSPNP-LP-----QGVPPSPGPPPLPSTTA-----PTLGGSGAPP---PPM 116

QY 178 KAVPTVAGARPLFHFKEPSPHPAPSPAGGHLGYDPTAAALSLPLGAAAAGSQA 237

DB 117 P--PPPLGSPFVIVSSMGSGPLPPAPPSPGVPSSPQINSTVSLPGGSG----- 166

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTSTASLLGESPLSPSPSRSSSGEGTCVAGCD 297

DB 167 -----PPEDVKP-----PVLGVRGLHCPPPGPGGAGKRLCAICGD 202

QY 298 NAAOCHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRCCYCRFOKCLSVGM 357

DB 203 RSSGKHGYVSGCEGCKGFFKRTIRKDLTYSRDNKCDTVDKQRNRCCYCRYKCLATGM 262

QY 358 KVEVVRTDSLKGR-----RGRLPSPKSPLOQEPSOPSPSPPICMNVALRALTD- 408

DB 263 KREAVQERQRKDXGDGECAGGAPEE-----MPVDRILEAEALVEQKSDQ 309

QY 409 -----STPRDLDSRYCPTDQAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIP 456

DB 310 GVEGPGGTGGSGSPND-----PVTNICOAAD-----KQLFTLV-----EWAKRIP 350

QY 457 GTPTDKEDOTLLIESAFLEFLVRLSIRSNATAEDKFFVFCNGLVLRHLCQ-LRGGEWLD 515

DB 351 HESSLPDDQVLLRAGMELLIAFSRHSIDVRGILLATLGLVHRNSAHSAGVAIFD 410

QY 516 ST-KDPSLNLQSLNDIOALACLALSMT-ERHGLKEPKRVEELCNKITSLSKDH-QSK 572

DB 411 RVLTSLVSKQRDRMDKTELGLCLRAILLNPDAKGLSNPSEVEVLRKRYASLETYCKQK 470

QY 573 GOALEFTESKVLGALVELRKICTLGLQRIFYKL 606

Db 471 YPEQGRFAKILLRLPALRSIGLKCLEHLEFFFKL 504

RESULT 14

US-08-216-592A-4

; Sequence 4, Application US/08216592A

; Patent No. 6635429

; GENERAL INFORMATION:

; APPLICANT: LEID, MARK

; APPLICANT: KASTNER, PHILIPPE

; APPLICANT: CHAMON, PIERRE

; TITLE OF INVENTION: NOVEL HETERODIMERIC STEROID RECEPTOR

; TITLE OF INVENTION: PROTEINS, GENES ENCODING SAME, AND USAGE THEREOF

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox

; STREET: 1100 New York Avenue NW Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/216,592A

; FILING DATE: 23-MAR-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/825,667

; FILING DATE: 24-JAN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: GOLDSTEIN, JORGE A

; REGISTRATION NUMBER: 29,021

; REFERENCE/DOCKET NUMBER: 1383.0060002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 533 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-216-592A-4

Query Match 15.9%; Score 530.5; DB 4; Length 533;

Best Local Similarity 29.5%; Pred. No. 1.2e-32;

Matches 150; Conservative 71; Mismatches 189; Indels 99; Gaps 17;

QY 118 PPASPEDEVLPSTSMYKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGGLDPPM 177

DB 83 PDSSSPNP-LP-----QGVPPSPGPPPLPSTTA-----PTLGGSGAPP---PPM 124

QY 178 KAVPTVAGARPLFHFKEPSPHPAPSPAGGHLGYDPTAAALSLPLGAAAAGSQA 237

DB 125 P--PPPLGSPFVIVSSMGSGPLPPAPPSPGVPSSPQINSTVSLPGGSG----- 174

QY 238 LESHYPGLPLAKRAAPLAPPLGLTPTSTASLLGESPLSPSPSRSSSGEGTCVAGCD 297

DB 175 -----PPEDVKP-----PVLGVRGLHCPPPGPGGAGKRLCAICGD 210

QY 298 NAAOCHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNRCCYCRFOKCLSVGM 357

DB 211 RSSGKHGYVSGCEGCKGFFKRTIRKDLTYSRDNKCDTVDKQRNRCCYCRYKCLATGM 270

QY 358 KVEVVRTDSLKGRRLSPKSPLOQEPSOPSPSPPICMNVALRAL----- 406

DB 271 KREAVQERQRG-----KQKDG-DGEGAGGAPEEMPVDRILEAEALVEQKSDQVEGP 322

QY 407 -----TDSTPRDLDSRYCPTDQAAGTDAEHVQOQFYNNLTASIDVSRSWAEKI PGFTDL 461

Db 323 GGTGGSGSPND-----PVNICOAD-----KQLFTLV-----EWAKRIHFSSSL 363  
Qy 462 PKEDQTLIESAFLELFLVLRISNTAEDKVFVFNGLVLRHLC- LRFGFWLDSI-KD 519  
Db 364 PLDDQVILLRACWNELLIASFSHRSDVRDGIILLATGLHVRHNSAHSAGVGAIFDRVLTE 423  
Qy 520 FSLNLSQSLNLDIQALACLALSMT- ERHGLKEPKRVEELCNKITSSILKDH-QSKGQALE 577  
Db 424 LVSKWRDMRMDKTELGCURLAILFPNDKAGLNPSEVEVLRKRYASLETYCKQKYPBOQ 483  
Qy 578 PTESKVLGALVELRKTCTGLQRIFYKL 606  
Db 484 GRFAKLLRLPALRSIGLKLHLEHFFFKL 512

RESULT 15  
US-09-949-016-11572  
; Sequence 11572, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11572  
; LENGTH: 577  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11572

Query Match 15.9%; Score 530.5; DB 4; Length 577;  
Best Local Similarity 29.4%; Pred. No. 1.3e-32;  
Matches 151; Conservative 70; Mismatches 189; Indels 103; Gaps 17;  
Qy 118 PPASSPEDEVLPSTMYFKQSPSTPTTTPAFPQAGALWDEALPSAPGCIAPGILLDPPM 177  
Db 123 PDSSSPNP--LP-----QGVPPPPPGGPPLPSTA-----PSLGGSGAPP--PPM 164  
Qy 178 KAVTVAGARFPLFHKPSPPHPPAPSPAGHHGLGYDPTAAALSLPLGAAAAAGSQA 237  
Db 165 P--PPPLGSPFVSSSMGSPGLPPPPAPPFGSPVSSPQINSTVSLPGGSG-- 214  
Qy 238 LESHVYGLPLAKRAAPLAPPPLGLTPSPASSLGLGESPLSPSPSSSGSGTCVCGD 297  
Db 215 -----PPEDVXP-----PVLGVRGLHCPPPPGPGGAGKRLCAICGD 250  
Qy 298 NAAOCHYGVRTCEGCKGFKRTVQKNAKYVCLANKNCVPDKRRNRQYCRFQKCLSVGM 357  
Db 251 RSSGKHGYVYSCGCKGFKRTIKDLTYSCRDNDKCTVDKEQRNRCQYCRYQKCLATGM 310  
Qy 358 VKEVVRTSLKGRRLSPKPSLPQEPSPQSPSPPICMNVALVRAL----- 406  
Db 311 KREAVQEEQRQC-----KDKDG-DGEGAGGAPPEMPVDRILEAEAVEQKSDQVEGP 362  
Qy 407 -----TDSFREDLDSRYCPTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDL 461  
Db 363 GGTGGSGSPND-----PVNICOAD-----KQLFTLV-----EWAKRIHFSSSL 403  
Qy 462 PKEDQTLIESAFLELFLVLRISNTAEDKVFVFNGLVLRHLC- LRFGFWLDSI-KD 516  
Db 404 PLDDQVILLRACWNELLIASFSHRSDVRDGIILLATGLHVRHNSAHSAGVGAIFDRSLR 463

Qy 517 -IKDFSLNLSQSLNLDIQALACLALSMT- BRHGLKEPKRVEELCNKITSSILKDH-QSKG 573  
Db 464 VTELVSXWRDMRMDKTELGCURLAILFPNDKAGLNPSEVEVLRKRYASLETYCKQKY 523  
Qy 574 QALEPTESKVLGALVELRKTCTGLQRIFYKL 606  
Db 524 PEQOGRFAKLLRLPALRSIGLKLHLEHFFFKL 556  
Search completed: March 12, 2005, 17:52:49  
Job time : 45 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Query			Description	
	Score	Match	Length	ID	
1	3133	99.9	626	9	US-09-853-386-64
2	3333	99.9	626	15	US-10-414-080-14
3	3316	99.4	626	9	US-09-853-386-65
4	3316	99.4	626	9	US-09-853-386-96
5	3316	99.4	626	17	US-10-659-004-116
6	3299.5	98.9	625	9	US-09-853-386-63
7	3299.5	98.9	625	15	US-10-414-080-13
8	3133.5	93.9	643	9	US-09-853-386-70
9	3133.5	93.9	643	15	US-10-414-080-17
10	3123	93.6	587	16	US-10-755-889-64
11	3074	92.1	628	9	US-09-853-386-66
12	3074	92.1	628	9	US-09-853-386-68
13	3074	92.1	628	9	US-09-853-386-73

```
QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
Db 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
QY 181 PTVAGARPLFHFKPSPPHPAPSPAGGHHILGYDPTAAALSLPLGAAAAAGSQAALLES 240
Db 181 PTVAGARPLFHFKPSPPHPAPSPAGGHHILGYDPTAAALSLPLGAAAAAGSQAALLEG 240
QY 241 HPYGLPLAKRAAPLAFPLPLGLTPSTASSLLGESPSLPSPPSSSGEGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAFPLPLGLTPSTASSLLGESPSLPSPPSSSGEGTCAVCGDNAA 300
QY 301 COHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNRQYCRFKCLSVGMVKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNRQYCRFKCLSVGMVKE 360
QY 361 VVRTDSLKGRGRRLSPKPSPLQOEPSQSPSPSPPICMNALVRALTDSTPRDLDSRYC 420
Db 361 VVRTDSLKGRGRRLSPKPSPLQOEPSQSPSPSPPICMNALVRALTDSTPRDLDSRYC 420
QY 421 PTDOAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480
Db 421 PTDOAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480
QY 481 RLSIRSTAEDKVFVFCNGLVHLRLQCLRGFGWLDSIKDPSLNQSLNLDIQALACLAL 540
Db 481 RLSIRSTAEDKVFVFCNGLVHLRLQCLRGFGWLDSIKDPSLNQSLNLDIQALACLAL 540
QY 541 SMITERHGLKEPKRVEELCNKITSSLDKHQSGQALEPTESKVLGALVELAKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLDKHQSGQALEPTESKVLGALVELAKICTLGLQR 600
QY 601 IFYKLKEDLVSPSSIIDKLFDLTLPF 626
Db 601 IFYKLKEDLVSPSSIIDKLFDLTLPF 626
```

## RESULT 2

```
US-10-414-080-14
; Sequence 14, Application US/10414080
; Publication No. US2003022088A1
; GENERAL INFORMATION:
; APPLICANT: MULLICAN, SHANNON E.
; APPLICANT: CONNEELY, ORLA M.
; APPLICANT: MILBRANDT, JEFFREY
; TITLE OF INVENTION: NOR-1 AND NUR77 NUCLEAR RECEPTORS AS TARGETS FOR
; FILE REFERENCE: P02454US1
; CURRENT APPLICATION NUMBER: US/10/414,080
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/373,238
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-080-14
```

Query Match 99.9%; Score 3333; DB 15; Length 626;  
Best Local Similarity 99.8%; Pred. No. 3.5e-217;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MPCVQAQYSPSPGSSYAAQTSYSEYTTIIMNPDYTKLTMDLGSTTEITATATSLPSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQTSYSEYTTIIMNPDYTKLTMDLGSTTEITATATSLPSIST 60
QY 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
```

```
Db 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
QY 181 PTVAGARPLFHFKPSPPHPAPSPAGGHHILGYDPTAAALSLPLGAAAAAGSQAALLES 240
Db 181 PTVAGARPLFHFKPSPPHPAPSPAGGHHILGYDPTAAALSLPLGAAAAAGSQAALLEG 240
QY 241 HPYGLPLAKRAAPLAFPLPLGLTPSTASSLLGESPSLPSPPSSSGEGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAFPLPLGLTPSTASSLLGESPSLPSPPSSSGEGTCAVCGDNAA 300
QY 301 COHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNRQYCRFKCLSVGMVKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNRQYCRFKCLSVGMVKE 360
QY 361 VVRTDSLKGRGRRLSPKPSPLQOEPSQSPSPSPPICMNALVRALTDSTPRDLDSRYC 420
Db 361 VVRTDSLKGRGRRLSPKPSPLQOEPSQSPSPSPPICMNALVRALTDSTPRDLDSRYC 420
QY 421 PTDOAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480
Db 421 PTDOAAAGTDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480
QY 481 RLSIRSTAEDKVFVFCNGLVHLRLQCLRGFGWLDSIKDPSLNQSLNLDIQALACLAL 540
Db 481 RLSIRSTAEDKVFVFCNGLVHLRLQCLRGFGWLDSIKDPSLNQSLNLDIQALACLAL 540
QY 541 SMITERHGLKEPKRVEELCNKITSSLDKHQSGQALEPTESKVLGALVELAKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLDKHQSGQALEPTESKVLGALVELAKICTLGLQR 600
QY 601 IFYKLKEDLVSPSSIIDKLFDLTLPF 626
Db 601 IFYKLKEDLVSPSSIIDKLFDLTLPF 626
```

## RESULT 3

```
US-09-853-386-65
; Sequence 65, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Breenihan, Barry
; APPLICANT: CONNEELY, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 65
; LENGTH: 626
; TYPE: PRT
; ORGANISM: HUMAN
US-09-853-386-65
```

Query Match 99.4%; Score 3316; DB 9; Length 626;  
Best Local Similarity 99.4%; Pred. No. 5e-216;  
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MPCVQAQYSPSPGSSYAAQTSYSEYTTIIMNPDYTKLTMDLGSTTEITATATSLPSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQTSYSEYTTIIMNPDYTKLTMDLGSTTEITATATSLPSIST 60
QY 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLDPPMKAV 180
```



```
Db 121 SPSPEVLSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLLDPKAV 180
QY 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAALSLPLGAAAAAGSQAALLES 240
Db 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAALSLPLGAAAAAGSQAALLEG 240
QY 241 HPYGLPLAKRAAPLAPFPPLGLTPSPASSLLGESPSLPPSPSSSSGGGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAPFPPLGLTPSPASSLLGESPSLPPSPSSSSGGGTCAVCGDNAA 300
QY 301 CQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVPDKERRNRCOYCRFKQKCLSVGMVKE 360
Db 301 CQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVPDKERRNRCOYCRFKQKCLSVGMVKE 360
QY 361 VVRTSLKGRRLPSKPKSPLOQEPSQSPSPPICMNALVRALTDTSTPRDLDSRYC 420
Db 361 VVRTSLKGRRLPSKPKSPLOQEPSQSPSPPICMNALVRALTDTSTPRDLDSRYC 420
QY 421 PTDQAAAGTDAEHVOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLIESAFLEFVL 480
Db 421 PTDQAAAGTDAEHVOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLIESAFLEFVL 480
QY 481 RLSIRNTAEDKFVFCNGLVHLRLQCLRGFGWLDSIKDFSLNLSLNLDIOALACLAL 540
Db 481 RLSIRNTAEDKFVFCNGLVHLRLQCLRGFGWLDSIKDFSLNLSLNLDIOALACLAL 540
QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKICTLGLQR 600
QY 601 IFYKLEDLVSPSSIIDKFLDTLPF 626
Db 601 IFYKLEDLVSPSSIIDKFLDTLPF 626

RESULT 4
US-09-853-386-96
; Sequence 96, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnahan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 96
; LENGTH: 626
; TYPE: PRT
; ORGANISM: HUMAN
US-09-853-386-96

Query Match 99.4%; Score 3316; DB 9; Length 626;
Best Local Similarity 99.4%; Pred. No. 5e-216; 3; Indels 0; Gaps 0;
Matches 622; Conservative 1; Mismatches 3;

QY 1 MFCVQAQYSPSPGSSYAAQYSSSEYTTIMPNPDYTKLTWDLGSTEITATATSLPSIST 60
Db 1 MFCVQAQYSPSPGSSYAAQYSSSEYTTIMPNPDYTKLTWDLGSTEITATATSLPSIST 60
QY 61 FVEGYSSNVELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNVELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHHHH 120
QY 121 SPSPEVLSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLLDPKAV 180
```

```
Db 121 SPSPEVLSTSMYFKQSPSPSTPTTAPFPQAGALWDEALPSAPGCIAPGLLDPKAV 180
QY 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAALSLPLGAAAAAGSQAALLES 240
Db 181 PTVAGARFPLFHFKSPSPHPAPSPAGGHHGLGYDPTAAALSLPLGAAAAAGSQAALLEG 240
QY 241 HPYGLPLAKRAAPLAPFPPLGLTPSPASSLLGESPSLPPSPSSSSGGGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAPFPPLGLTPSPASSLLGESPSLPPSPSSSSGGGTCAVCGDNAA 300
QY 301 CQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVPDKERRNRCOYCRFKQKCLSVGMVKE 360
Db 301 CQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVPDKERRNRCOYCRFKQKCLSVGMVKE 360
QY 361 VVRTSLKGRRLPSKPKSPLOQEPSQSPSPPICMNALVRALTDTSTPRDLDSRYC 420
Db 361 VVRTSLKGRRLPSKPKSPLOQEPSQSPSPPICMNALVRALTDTSTPRDLDSRYC 420
QY 421 PTDQAAAGTDAEHVOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLIESAFLEFVL 480
Db 421 PTDQAAAGTDAEHVOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDOTLIESAFLEFVL 480
QY 481 RLSIRNTAEDKFVFCNGLVHLRLQCLRGFGWLDSIKDFSLNLSLNLDIOALACLAL 540
Db 481 RLSIRNTAEDKFVFCNGLVHLRLQCLRGFGWLDSIKDFSLNLSLNLDIOALACLAL 540
QY 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTESKVLGALVELRKICTLGLQR 600
QY 601 IFYKLEDLVSPSSIIDKFLDTLPF 626
Db 601 IFYKLEDLVSPSSIIDKFLDTLPF 626

RESULT 5
US-10-659-004-116
; Sequence 116, Application US/10659004
; Publication No. US20050048507A1
; GENERAL INFORMATION:
; APPLICANT: Zhong et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
; FILE REFERENCE: 21402-608
; CURRENT APPLICATION NUMBER: US/10/659,004
; CURRENT FILING DATE: 2003-09-09
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298,285
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/298,556
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/299,949
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 60/300,883
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 116
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-659-004-116

```
Query Match          99.4%; Score 3316; DB 17; Length 626;
Best Local Similarity 99.4%; Pred. No. 5e-216;
Matches 622; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MPCVQAQYSPSPGSSYAAQYSSYTTIINNPDYTKLTMDLGSFEITATATTSLSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQYSSYTTIINNPDYTKLTMDLGSFEITATATTSLSIST 60

Qy 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

Qy 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPPMKAV 180
Db 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPPMKAV 180

Qy 181 PTVAGARPLFHFKPSPPHPPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALES 240
Db 181 PTVAGARPLFHFKPSPPHPPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALEG 240

Qy 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLIGESPSLPPSSSSGEGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLIGESPSLPPSSSSGEGTCAVCGDNAA 300

Qy 301 COHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSYGMVKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSYGMVKE 360

Qy 361 VVRTSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNALVRALTSTPRDLDSRYC 420
Db 361 VVRTSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNALVRALTSTPRDLDSRYC 420

Qy 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480

Qy 481 RLSIRSTAEDKVFVNCGLVLRQLCLRGFGWLDISKDFSLNQLSLNLDIQALACLAL 540
Db 481 RLSIRSTAEDKVFVNCGLVLRQLCLRGFGWLDISKDFSLNQLSLNLDIQALACLAL 540

Qy 541 SMITERHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600

Qy 601 IFYKLEDLVSPSSIIDKFLDLP 626
Db 601 IFYKLEDLVSPSSIIDKFLDLP 626
```

RESULT 6

```
US-09-853-386-63
; Sequence 63, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnahan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63
; LENGTH: 625
; TYPE: PRT
; ORGANISM: HUMAN
```

US-09-853-386-63

```
Query Match          98.9%; Score 3299.5; DB 9; Length 625;
Best Local Similarity 99.2%; Pred. No. 6.5e-215;
Matches 621; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 MPCVQAQYSPSPGSSYAAQYSSYTTIINNPDYTKLTMDLGSFEITATATTSLSIST 60
Db 1 MPCVQAQYSPSPGSSYAAQYSSYTTIINNPDYTKLTMDLGSFEITATATTSLSIST 60

Qy 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120
Db 61 FVEGYSSNYELKPCSVYQMRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHHHHHH 120

Qy 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPPMKAV 180
Db 121 SSPDEVLPSTSMYFKQSPPTPTTAPPPQAGALWDEALPSAPGCIAPGLLDPPMKAV 180

Qy 181 PTVAGARPLFHFKPSPPHPPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALES 240
Db 181 PTVAGARPLFHFKPSPPHPPAPSPAGGHHLYDPTAAALSLPLGAAAAAGSQAALEG 240

Qy 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLIGESPSLPPSSSSGEGTCAVCGDNAA 300
Db 241 HPYGLPLAKRAAPLAPPLGLTPSTASSLIGESPSLPPSSSSGEGTCAVCGDNAA 300

Qy 301 COHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSYGMVKE 360
Db 301 COHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRRCQYCRFKCLSYGMVKE 360

Qy 361 VVRTSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNALVRALTSTPRDLDSRYC 420
Db 361 VVRTSLKGRGRPLSPKPSPLQOEPSQSPSPSPPICMNALVRALTSTPRDLDSRYC 420

Qy 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480
Db 421 PTDOAAAGTDAHVQOOFYNLLTASIDVSRSWAEKIPGFTDLPKEDOTLLIESAFLELVL 480

Qy 481 RLSIRSTAEDKVFVNCGLVLRQLCLRGFGWLDISKDFSLNQLSLNLDIQALACLAL 540
Db 481 RLSIRSTAEDKVFVNCGLVLRQLCLRGFGWLDISKDFSLNQLSLNLDIQALACLAL 540

Qy 541 SMITERHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600
Db 541 SMITERHGLKEPKRVEELCNKITSSLKHQSKGQALEPTESKVLGALVELRKICTLGLQR 600

Qy 601 IFYKLEDLVSPSSIIDKFLDLP 626
Db 601 IFYKLEDLVSPSSIIDKFLDLP 626
```

RESULT 7

```
US-10-414-080-13
; Sequence 13, Application US/10414080
; Publication No. US2003020288A1
; GENERAL INFORMATION:
; APPLICANT: MULLICAN, SHANNON E.
; APPLICANT: CONNEELY, ORLA M.
; APPLICANT: MILBRANDT, JEFFREY
; TITLE OF INVENTION: NOR-1 AND NUR77 NUCLEAR RECEPTORS AS TARGETS FOR
; FILE REFERENCE: P02454US1
; CURRENT FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/373,238
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-414-080-13
```













**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3299.5	98.9	625	2	S71930	neuron-derived rec
2	3074	92.1	628	2	JC2493	neuron-derived op
3	2032	60.9	430	2	S66671	neuron-derived rec
4	1751	52.5	598	2	A57040	f-cell nuclear rec
5	1746	52.3	598	2	A46225	nuclear orphan rec
6	1728	51.8	598	2	I84692	nuclear orphan h
7	1367	41.0	598	2	A37251	probable nuclear h
8	1350.5	40.5	577	2	S33763	hormone receptor N
9	1116.5	39.5	601	1	QWMSN1	probable hormone r
10	1277	38.3	563	2	JQ0623	nerve growth facto
11	1167	35.0	551	2	S58238	DHR38 protein - fr
12	1023	30.7	360	2	S58205	DHR38 protein - si
13	554.5	16.6	619	1	T43348	nuclear steroid ho
14	554	16.6	562	2	G88408	protein nhr-6 [imp
15	543	16.3	455	2	S13512	retinoic acid rece
16	531.5	15.9	462	2	S05050	retinoic acid nucl
17	530.5	15.9	462	1	A29491	retinoic acid rece
18	529.5	15.9	533	2	S37781	retinoid X recepto
19	525.5	15.7	446	2	A34418	H-2 region II bind
20	521.5	15.6	448	2	D41727	retinoid X recepto
21	521.5	15.6	520	2	I84718	RXR-beta1 isoform
22	521	15.6	955	4	C40045	probable transcrip
23	520.5	15.6	458	2	S06123	retinoic acid rece
24	519.5	15.6	451	2	A41651	retinoic acid rece
25	514.5	15.4	459	2	A41977	retinoic acid rece
26	514.5	15.4	476	2	B43977	retinoic acid rece
27	511.5	15.3	448	2	A43786	retinoic acid rece
28	509	15.3	441	2	I50515	retinoid X recepto
29	507	15.2	448	2	S05051	retinoic acid rece



```
QY 117 IPPASSPEDEVLPSTSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGILLDDPP 176
Db 119 IPPSGPEDEVLPSTSMYFKQSPSPSTPTTGGPPQAGALWDELPSAPGCIAPGILLDDPPQ 178
QY 177 MKAVPTV-AGARFPLFHFKSPSPHPAPSPAGGHHLYGSDPTAAALSLPLGAAAAAGSQA 235
Db 179 MKAVPPMAAAARFPIF-FKPSPPHPAPSPAGGHHLYGSDPTAAALSLPLGAAAAAGSQA 237
QY 236 AALESHPYGLPLAKRAAPLAPFPLGLTPSPPTASSLLGESPSLPSPPSRSSSGEGTCVAC 295
Db 238 AALEGHPYGLPLAKRTATLTFPLGLTASPTASSLLGESPSLPSPPNRRSSSGEGTCVAC 297
QY 296 GDNAAACOHYGVRTCECKGFFKRTVOKNAKYVCLANKNCVDPKRRNRQCYCRFOKCLSV 355
Db 298 GDNAAACOHYGVRTCECKGFFKRTVOKNAKYVCLANKNCVDPKRRNRQCYCRFOKCLSV 357
QY 356 GMVKEVVRTDSLKGRGRRLPSKPSLQOQSPSPSPSPPICMNALVRALTDSTPRDLJ 415
Db 358 GMVKEVVRTDSLKGRGRRLPSKPSLQOQSPSPSPSPPICMNALVRALTDSTPRDLJ 417
QY 416 YSR 418
Db 418 YSR 420

RESULT 4
A57040
T-cell nuclear receptor NOT - human
C:Species: Homo sapiens (man)
C>Date: 03-Oct-1995 #sequence revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A57040; I56260; S38636
R:Mages, H.W.; Rilke, O.; Bravo, R.; Senger, G.; Krocze, R.A.
Mol. Endocrinol. 8, 1583-1591, 1994
A:Title: NOT, a human immediate-early response gene closely related to the steroid/thyroid
A:Reference number: A57040; MUID:95183071; PMID:7877627
A:Accession: A57040
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-598 <NAG>
A:Cross-references: UNIPROT:P43354; EMBL:X75918; NID:G415822; PIDN:CAA53518.1; PID:G4158
R:Okabe, T.; Takayanagi, R.; Imasaki, K.; Hajji, M.; Nawata, H.; Watanabe, T.
J. Immunol. 154, 3871-3879, 1995
A:Title: cDNA cloning of a NGFI-B/nur77-related transcription factor from an apoptotic h
A:Reference number: I56260; MUID:95221892; PMID:7706727
A:Accession: I56260
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 64-464, 'W', 466-467, 'W', 469-598 <RES>
A:Cross-references: GB:S77154; NID:G913966; PIDN:AAB33999.1; PID:G913967
C:Genetics:
A:Gene: GDB:NURR1; TINUR; NOT
A:Cross-references: GDB:624567; OMIM:601828
A:Map position: 2q22-2q23
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; zinc finger
F:261-514/Domain: erba transforming protein homology <ERBA>

Query Match 52.5%; Score 1751; DB 2; Length 598;
Best Local Similarity 54.5%; Pred. No. 4.7e-89;
Matches 363; Conservative 74; Mismatches 121; Indels 108; Gaps 18;
QY 1 MPCVOAQYSPSPGSSYAAQTY-----SSYTTTEIMNPDYTKLTMDLGSTEITATATSLP 56
Db 1 MPCVOAQYSSPQGASPASQSYSHSGEYSSDFLTPEFVKFMDLTNTEI--TATTSLP 58
QY 57 S1STFVEGYSSNYELKPSVCYQM-----QRPLKVEGRAPSYHHHHHHHHHHHHHHHHQOQH 112
Db 59 SFSTFMDNYSTGYDVKPCLYQMPLSGQSSIKVEDIQMNTQQHSH----- 105
QY 113 QOPSIPPASSPEDEVLP-STSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGP 171
Db 106 ----LPPQS---EEMPHSGSVYKPSPTTTTCFQVQHSPPMDD-----PGS 148
```

```
QY 172 LLDPPMKAVPT-----VAGARFPLFHFKSPSPHPA-----PSPAG 207
Db 149 LNFHONYVATTHMIEQRTKTPVSRSLFSFKOSPPGTPVSSQMRPDGDLPHVPMNPEPAG 208
QY 208 GHHLGYDPTAAALSLPLGAAAAAGSQAALSHSPYGLP-LAKRAAPLAPFPLGLTPSP 266
Db 209 SHV-----VDQITFAVENPIRKPSMGFPGLQI---GH 239
QY 267 ASLLGESPSLPSPPSRSSSGEGTCVACGDNAAACOHYGVRTCECKGFFKRTVOKNAKY 326
Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACOHYGVRTCECKGFFKRTVOKNAKY 297
QY 327 VCLANKNCVDPKRRNRQCYCRFOKCLSVGMVKEVVRTDSLKGRGRRLPSKPSLQOQEP 386
Db 298 VCLANKNCVDPKRRNRQCYCRFOKCLAVGMVKEVVRTDSLKGRGRRLPSKPS-----P 352
QY 387 SOPSPPSPPICMNALVRALTDSTP--RDLDSRY-CPTDQAAAAGTDAEHVQOYFVNLTA 443
Db 353 QEPSPPSPVSLISALVRAHVDSNPAMTSLDYSRQANPDYQMSGDDTQHIQQFYDLLTG 412
QY 444 SIDVRSWAEEKIPGFTDLPKEDQTLIESAFLEFVLRISIRSNATAEDKPFVFCNGLVLRH 503
Db 413 SMEIIRGWAEEKIPGADLPKADQDLIFESAFLEFVLRILAYRSNPVEGKLIFCNGVVLHR 472
QY 504 LQCLRGCEWLDLSIKDPSNLQSLNLDIOALCLALSMTTERHGLKEPKRVEELCNKIT 563
Db 473 LQCVRGGEWIDSIVEFSSNLQNMIDISAFSIAALAMVTERHGLKEPKRVEELQNKIV 532
QY 564 SSLKDHQ--SKQALPEPE-SKVLGALVELRIKTLGLQRIFYKLKLELVSPSSIDKLF 620
Db 533 NCLKDHVTFNNGGLNRPVLSKLLGLPELRLCTQGLQRIFYKLKLELVPPPAIDKLF 592
QY 621 LDTLPP 626
Db 593 LDTLPP 598

RESULT 5
A46225
nuclear orphan receptor Nurrl - mouse
N:Alternate names: Nur-related transcription factor 1
C:Species: Mus musculus (house mouse)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A46225
R:Law, S.W.; Conneely, O.M.; DeMayo, F.J.; O'Malley, B.W.
Mol. Endocrinol. 6, 2129-2135, 1992
A:Title: Identification of a new brain-specific transcription factor, NURR1.
A:Reference number: A46225; MUID:93149122; PMID:1491694
A:Accession: A46225
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-598 <LAW>
A:Cross-references: UNIPROT:Q06219; GB:S53744; NID:G264410; PIDN:AAB25138.1; PID:G26441
A:Experimental source: brain
A>Note: sequence extracted from NCBI backbone (NCBI:123734, NCBI:123736)
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger
F:261-514/Domain: erba transforming protein homology <ERBA>

Query Match 52.3%; Score 1746; DB 2; Length 598;
Best Local Similarity 54.4%; Pred. No. 8.8e-89;
Matches 362; Conservative 75; Mismatches 121; Indels 108; Gaps 18;
QY 1 MPCVOAQYSPSPGSSYAAQTY-----SSYTTTEIMNPDYTKLTMDLGSTEITATATSLP 56
Db 1 MPCVOAQYSSPQGASPASQSYSHSGEYSSDFLTPEFVKFMDLTNTEI--TATTSLP 58
QY 57 S1STFVEGYSSNYELKPSVCYQM-----QRPLKVEGRAPSYHHHHHHHHHHHHHHHHQOQH 112
Db 59 SFSTFMDNYSTGYDVKPCLYQMPLSGQSSIKVEDIQMNTQQHSH----- 105
QY 113 QOPSIPPASSPEDEVLP-STSMYFKQSPSPSTPTTAPPPQAGALWDEALPSAPGCIAPGP 171
Db 106 ----LPPQS---EEMPHSGSVYKPSPTTTTCFQVQHSPPMDD-----PGS 148
```

Db 106 ----LPPQS---EEMPHSGSVYKPPSPPTSTPSTFQVQHSPPWDD-----PGS 148

Qy 172 LDDPKMAVPT-----VAGARPLFHFKPSPPHPEA-----PSPAG 207

Db 149 LHNHFNQVAVTHMIEQRKTPVSRSLSLFSFKQSPGTEVSSCOMRFDGFLHVPNMNPEPAG 208

Qy 208 GHHLGYDPTAAALSLPLGMAAAGSQAALASHPHYGLP-LAKRAAPLAFPLGLTSPPT 266

Db 209 SHV-----VDGQTFVAVPNPIRKPASMGFFGLQI---GH 239

Qy 267 ASSLLGESPSLPPSSSSSGEGTCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNKY 326

Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNKY 297

Qy 327 VCLANKNCVDKRRNRNCOYCRFOKCLSVGMVKEVVRTDSLKGRGRGLPSKPKSLQOEP 386

Db 298 VCLANKNCVDKRRNRNCOYCRFOKCLAVGMVKEVVRTDSLKGRGRGLPSKPKS-----P 352

Qy 387 SOPSPPSPPICMNALVRALTDTSP--RDLDSRY--CPTDQAAAGTDAEHVQOQFYNLLTA 443

Db 353 QDPSPPSPVSLISALVRAHVDSNPAMTSLDYSRFOANPDYQMSGDDTQHIOQFYDLLTG 412

Qy 444 SIDVRSWAEEKIPGFTDLPKEDQTLLESFALEFVLRLSIRNSTAEDKFFVFCNGVLHR 503

Db 413 SMEIIRGWAEEKIPGFADLPKADQDLLESFALEFVLRLAYRSNPVEGKLIFCNGVLHR 472

Qy 504 LQCLRGFGWLDISIKDFSLNLSQSLNDIOALACLALSALSMITERHGLKPKRVEELCNKIT 563

Db 473 LQCVRGFGWIDSIIVEFSSNLQNMNIDISAFSCIAALAMVTERHGLKPKRVEELCNKIV 532

Qy 564 SSKLDHQ--SKGQALEPTE--SKVLGALVELRKTGLQRIFFYLKLEDLVSPPSIIDKLF 620

Db 533 NCLKDHVTFNNGGLNRPNYLSKLLGKLPRLTCTOGLQRIFFYLKLEDLVPPPAIIDKLF 592

Qy 621 LDTLPF 626

Db 593 LDTLPF 598

RESULT 6.

184692

nuclear orphan receptor HZF-3 - rat

N;Alternate names: thyroid/steroid receptor homolog RNR-1

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 09-Jul-2004

C;Accession: I84692; A46050

R;Pena de Ortiz, S.; Camon, M.M.; Jamieson, G.A.

A;Title: Expression of nuclear hormone receptors within the rat hippocampus: identification

A;Reference number: I48178; MUID:94335560; PMID:7914660

A;Accession: I84692

A;Status: preliminary; translated from GB/EMBL/DBU

A;Molecule type: mRNA

A;Residues: 1-598 <PEN>

A;Cross-references: UNIPROT:Q07917; EMBL:U01146; NID:g515968; PIDN:AAC52143.1; PID:g5415

R;Seearce, L.M.; Laz, T.M.; Hazel, T.G.; Lau, L.F.; Taub, R.

J. Biol. Chem. 268, 8855-8861, 1993

A;Title: RNR-1, a nuclear receptor in the NGFI-B/Nur77 family that is rapidly induced in

A;Reference number: A46050; MUID:93232045; PMID:8473329

A;Accession: A46050

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-181, 'A', 183-249, 'PSP', 253-361, 'G', 363, 'DQCPRQT', 372-433, 'S', 435-598 <SCB>

A;Cross-references: GB:L08595; NID:g310215; PIDN:AAA42058.1; PID:g310216

A;Experimental source: liver and brain

A;Note: sequence extracted from NCBI backbone (NCBIN:129782, NCBIPI:129783)

C;Genetics:

A;Gene: HZF-3

C;Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog

C;Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; z

F;261-514/Domain: erba transforming protein homology <ERBA>

Query Match 51.8%; Score 1728; DB 2; Length 598;

Best Local Similarity 54.1%; Pred. No. 8.6e-88;

Matches 360; Conservative 74; Mismatches 124; Indels 108; Gaps 18;

Qy 1 MPCVQAGSPSPGSSVAAQTY---SSEYTTIMAPDYTKLTWDLGSTEITATATSLP 56

Db 1 MPCVQAGSPGSPGSSVAAQTY---SSEYTTIMAPDYTKLTWDLGSTEITATATSLP 58

Qy 57 S1STFVSGYSNTVELKPCSVYQM---QREPLIKVEGRAPSYHHHHHHHHHHHHHHHH 112

Db 59 SFSTFMDNYSTGYDVKPPCLYQMLSQQSSIKVEDIQMHNYQOHS- 105

Qy 113 QQPSIIPASSPDEVLVLP-STSMYFKQSPPTPTTAPPPQAGALWDEALPSAPCIIAPGP 171

Db 106 ----LPPQS---EEMPHSGSVYKPPSPPTSTPSTFQVQHSPPWDD-----PGS 148

Qy 172 LDDPKMAVPT-----VAGARPLFHFKPSPPHPEA-----PSPAG 207

Db 149 LHNHFNQVAVTHMIEQRKTPVSRSLSLFSFKQSPGTEVSSCOMRFDGFLHVPNMNPEPAG 208

Qy 208 GHHLGYDPTAAALSLPLGMAAAGSQAALASHPHYGLP-LAKRAAPLAFPLGLTSPPT 266

Db 209 SHV-----VDGQTFVAVPNPIRKPASMGFFGLQI---GH 239

Qy 267 ASSLLGESPSLPPSSSSSGEGTCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNKY 326

Db 240 ASQLL--DTQVPSPPSRGSPSNEGLCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNKY 297

Qy 327 VCLANKNCVDKRRNRNCOYCRFOKCLSVGMVKEVVRTDSLKGRGRGLPSKPKSLQOEP 386

Db 298 VCLANKNCVDKRRNRNCOYCRFOKCLAVGMVKEVVRTDSLKGRGRGLPSKPKS-----P 352

Qy 387 SOPSPPSPPICMNALVRALTDTSP--RDLDSRY--CPTDQAAAGTDAEHVQOQFYNLLTA 443

Db 353 QDPSPPSPVSLISALVRAHVDSNPAMTSLDYSRFOANPDYQMSGDDTQHIOQFYDLLTG 412

Qy 444 SIDVRSWAEEKIPGFTDLPKEDQTLLESFALEFVLRLSIRNSTAEDKFFVFCNGVLHR 503

Db 413 SMEIIRGWAEEKIPGFADLPKADQDLLESFALEFVLRLAYRSNPVEGKLIFCNGVLHR 472

Qy 504 LQCLRGFGWLDISIKDFSLNLSQSLNDIOALACLALSALSMITERHGLKPKRVEELCNKIT 563

Db 473 LQCVRGFGWIDSIIVEFSSNLQNMNIDISAFSCIAALAMVTERHGLKPKRVEELCNKIV 532

Qy 564 SSKLDHQ--SKGQALEPTE--SKVLGALVELRKTGLQRIFFYLKLEDLVSPPSIIDKLF 620

Db 533 NCLKDHVTFNNGGLNRPNYLSKLLGKLPRLTCTOGLQRIFFYLKLEDLVPPPAIIDKLF 592

Qy 621 LDTLPF 626

Db 593 LDTLPF 598

RESULT 7

A37251

Probable nuclear hormone receptor NAK1 - human

N;Alternate names: early response protein NAK1; nuclear hormone receptor ST-59; steroid

C;Species: Homo sapiens (man)

C;Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 09-Jul-2004

C;Accession: A37251; A61191; A61451

R;Nakai, A.; Kartha, S.; Sakurai, A.; Toback, F.G.; DeGroot, L.J.

Mol. Endocrinol. 4, 1438-1443, 1990

A;Title: A human early response gene homologous to murine nur77 and rat NGFI-B, and rela

A;Reference number: A37251; MUID:91133413; PMID:2283997

A;Accession: A37251

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-598 <NAK>

A;Cross-references: UNIPROT:P22736; GB:D49728; NID:g1813881; PIDN:BAA08565.1; PID:g18138

R;Bondy, G.P.

Cell Growth Differ. 2, 203-208, 1991

A;Title: Phorbol ester, forskolin, and serum induction of a human colon nuclear hormone

A;Reference number: A61191; MUID:91329278; PMID:1651101

**RESULT 8**

S33763 hormone receptor NGFI-B homolog - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S33763; S31713  
F:Smith, T.S.; Matharu, P.J.; Sweetney, G.E.  
Biochim. Biophys. Acta 1173, 239-242, 1993  
A>Title: Cloning and sequencing of a Xenopus homologue of the inducible orphan receptor  
A:Reference number: S33763; MUID: 93277961; PMID: 8504173  
A:Accession: S33763  
A:Molecule type: mRNA  
A:Residues: 1-577 <SMI>  
A:Cross-references: UNIPROT:Q04913; EMBL:X70700; NID:g64916; PIDN:CAA50031.1; PID:g6491  
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolo  
C:Keywords: DNA binding; nucleus; transcription factor; zinc finger  
P;244-493/domain: erba transforming protein homology <ERBA>

	Query Match	40.5%	Score 1350.5	DB 2	Length 577
	Best Local Similarity	47.1%	Pred. No. 4.4e-67		
	Matches 305	Conservative 88	Mismatches 162	Indels 93	Gaps 19
Qy	1 MPCVQAQYSPSPGSGSVAATYSSEYTEIMNDPYTKLTMDLGSTEITATATSLPSIST	60			
Dd	1 MPCTIAQH-----GSLSQCACPCDNYVPDIINSEFGKFTMDLVNSEI--AASTSLSPFSST	53			
Qy	61 FVGEGSYNNELKPSCVVQM---ORPLIKVEEGRAPSYHHHHHHHHHHHHHQQOQQOPS	116			
Dd	54 PMDGYTGEPD---AFLYQIPSSNQSSLKVEE-----FOVFPG	87			
Qy	117 IPPAS--SPDEVELPSTMVFKQSPPS--TPTTPAPPPOAGALWDEALPSAPGCIAPGPL	172			
Dd	88 CYPGSFNTQLDETWSGSDGYGSPCSIIPSPTPGQNQLPTWE----CSYGAYSPQN	143			
Qy	173 LDP-----PMKAVPTVAGARPLLPHFKXPSPHPAPSPAGGHHLGYDPTAAAALSPLGAA	228			
Dd	144 YDNMRHWTEQQKNISQOTF--FSF-GTPAHSF-----NMAANPLKIATATHRLDQQ	192			
Qy	229 AAGSQAAALESHPHYGLPLAKRAAPLAFPPLGLTPTASSLLGESPLSPSPSRSSSG	288			
Dd	193 LVDTDFVALAQNSSAGFPFAV-----PLGOAPGVLDSSVLDSPLSPS-KTRSPSSN	242			
Qy	289 EGTCVAVCGDNAACHYGVRTCECKGFFKRTVKNAKYVCILANKNCVPDKRRNRRCQYCR	348			
Dd	243 EGRCAVCGDNASCQHYGVRTCECKGFFKRTVKNAKYICLANCKDCPDVKRRNRRCQFCR	302			
Qy	349 FQCLSVGMVKEVVRTDSLKGRRGRLPSPKPSLQOEPPSQSPSPPICMNALVRALTD	408			
Dd	303 FOXCLVVMGVKEVVRTDSLKGRRGRLPSPKPKQIAE-----SSPVDLINSIVRAHID	353			
Qy	409 STPRD--LDYSRYCPTDDAAAGTDAB-HVQOFYNLITASIDVRSWAEEKIPGTDLDPKD	465			
Dd	354 SIPESSKLDDYKPFQETVPLQLEKESVDVQQFYDLLSGSLDEVIRKWAEEKIQGFVDLPKD	413			
Qy	466 QTLIIESAFLFVLRLSIRSNTAEOKFPFCNGLVLRQLRGFGEWLDSIKDFSINLQ	525			
Dd	414 QDULLESAFLFVLRLVARSREBEKGLIFCNQGVULHRTQCVRGFEWDISIIIEFBSHSLQ	473			
Qy	526 SLNLIDIQAACLALSAMITERHGLKEPKRVVEELCNKITSLSKDH-----QSKGOALEP	578			
Dd	474 RMNIDVPSFCLSALVIVDRHGLKEPKRVVEELQSQI INCLKEHIPSSMNEQNPNCLE--	531			
Qy	579 TESKVIGALVELRKICTLGLQRIFYLKLEDVLPSPSIIDKLFIDLTLPP	626			
Dd	532 --SKLLGKLPELRTLCTQGLQRIFYLKLEDVLPSPSIIDKLFIDLTLPP	577			

**RESULT 9**

ORMSN1 probable hormone receptor N10, nuclear - mouse  
N:Alternate names: steroid hormone receptor homolog nur/77

RESULT 9  
QRMSN1  
Probable hormone receptor N10, nuclear - mouse  
N:Alternate names: steroid hormone receptor homolog nmr/77

C/Species: Mus musculus (mouse mouse)  
C/Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text change 09-Jul-2004

C:Accession: S06953; A30059; A31344  
R:Ryseck, R.P.; Macdonald-Bravo, H.; Mattei, M.G.; Ruppert, S.; Bravo, R.  
EMBO J. 8, 3327-3335, 1989  
A:Title: Structure, mapping and expression of a growth factor inducible gene encoding a  
A:Reference number: S06953; MUID:90059925; PMID:2555161  
A:Accession: S06953  
A:Molecule type: DNA  
A:Residues: 1-601 <RYS>  
A:CROSS-references: UNIPROT:P12813; EMBL:X16995; NID:G53312; PIDN:CAA34862.1; PID:G53313  
R:Hazel, T.G.; Nathans, D.; Lau, L.F.  
Proc. Natl. Acad. Sci. U.S.A. 85, 8444-8449, 1988  
A:Title: A gene inducible by serum growth factors encodes a member of the steroid and th  
A:Reference number: A94211; MUID:89042200; PMID:3186734  
A:Accession: A30059  
A:Molecule type: mRNA  
A:Residues: 1-601 <HAZ>  
A:CROSS-references: GB:J04113; NID:G200115; PIDN:AAA39843.1; PID:G200116  
C:Genetics:  
A:Map position: 15P  
A:Introns: 295/3; 339/1; 389/3; 457/2; 517/1  
C:Superfamily: probable hormone receptor N10, nuclear; erba transforming protein homolog  
C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription factor; zinc f  
F:268-517/Domain: erba transforming protein homology <ERBA>  
F:270-290/Region: zinc finger CCCC motif  
F:276-335/Domain: DNA binding #status predicted <DNA>  
F:306-330/Region: zinc finger CCCC motif

Query Match 39.5%; Score 1316.5; DB 1; Length 601;  
Best Local Similarity 44.9%; Pred. No. 3.4e-65;  
Matches 307; Conservative 78; Mismatches 158; Indels 141; Gaps 22;

QY 1 MPCVQAQY-----SPSPGSSYAAQYSEYITEIMNPDYTKLTMDLGSTTEITATATLSL 55  
DB 1 MFCIQAYGTATSPGPG-----RDHLTGDPALALEFGKPTWDLASPETAAPATL 50

QY 56 PSITFVEGYSNYELKPCVYQM---QRPLIKVEGRAPSYHHHHHHHHHHHHHHHHH 112  
DB 51 PSFSTFMDGYTGFEFD---TFLYQLFGTGTQPCSSACSASST-----SSS 91

QY 113 QQPSITPPASS-----PEDEVLPSTSMYFKQSPPS--TPTTAPFPQ 151  
DB 92 SSSATSPASASPKFEDFQVYGCYPTLSCGLDETSSSGSEYGGPCAPSPSTNFQPS 151

QY 152 AGALWD-----EALPSAPGCIAPGLLDPMPKAVPTVAGARFPLFH 192  
DB 152 QLSPWDGSGFHFSPQTYEGLVWTEQLPKA---SSGP---PP-----PPTF 192

QY 193 FKPSPPHPAPSPAGGHHLGVDPTAAALSPLGAAAAAGSQAALESHYPGLPLAKRA 252  
DB 193 FFSFSPPTGSPSLAQSLLKLPFPATHQLG-----EGESYSM----- 229

QY 253 PLAFPLGLT-PSPTASSLGSPLSPSPSSSSSG--EGTCVAGDAAACQHYGVRTC 309  
DB 230 PAAFPGLAPTSNRTSGIL-DAP-VTSKTSRSGASGGSEGCACVCGDNASCQHYGVRTC 287

QY 310 EGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFKCLSGVMYKEVVRTDSLKG 369  
DB 288 EGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFKCLAVGMVKEVVRTDSLKG 347

QY 370 RGRPLSPKPKPLQEPSPSPSPPICMNALVRALTDSTP--RDLDYSRYCPTDQAAA 427  
DB 348 RGRPLSPKPKQP-----PDASPTNLLTSILRAHLDSPGSKFQELVLPF 398

QY 428 G-TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESFAFELFVLSIRS 486  
DB 399 GKHEDAGDVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLESFAFELFVLSIRS 458

QY 487 NTAEDKVFVFCNGLVHLRLQCLRGFGWLDISIKDFSLNQLSLNDIQAALCSALSMTIR 546  
DB 459 KPGEGLKIFCSGLVLHQLQCARFGDWIDNILAFSRSLHSLGVDVPAPACLSALVLITDR 518

QY 547 HGLKEPKRVVELCNKLTSSLKDHQSKGQALEPTE-----SKVLGALVELRKICTLGLQRTF 602  
DB 547 HGLKEPKRVVELCNKLTSSLKDHQSKGQALEPTE-----SKVLGALVELRKICTLGLQRTF 602

DB 519 HGLQDPRRVELQNRIASCLAEHWAT-VAGDPPQASCLSRLLGKLPELRLCTQLQRI 577

QY 603 YLKLEDLVSPPSIIDKFLDLTLPF 626  
DB 578 CLKLEDLVPPPPVIVDKIFMDTLSP 601

RESULT 10  
JQ0623  
Nerve growth factor-inducible protein B - rat  
N:Alternate names: glucocorticoid receptor homolog  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
A:Accession: JQ0623; A34097  
R:Milbrandt, J.  
Neuron 1, 183-188, 1988  
A:Title: Nerve growth factor induces a gene homologous to the glucocorticoid receptor g  
A:Reference number: JQ0623; MUID:90166506; PMID:3272167  
A:Accession: JQ0623  
A:Molecule type: mRNA  
A:Residues: 1-563 <MIL>  
A:CROSS-references: GB:U17254; NID:G596053; PIDN:AAA56770.1; PID:G596054  
A:Experimental source: cell line PCL2  
R:Watson, M.A.; Milbrandt, J.  
Mol. Cell. Biol. 9, 4213-4219, 1989  
A:Title: The NGFI-B gene, a transcriptionally inducible member of the steroid receptor g  
A:Reference number: A34097; MUID:90066426; PMID:2479823  
A:Accession: A34097  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 255-260; 298-303; 349-354; 417-422; 476-481 <WAT>  
C:Comment: This protein is proposed to be a ligand-dependent transcriptional activator w  
F:230-479/Domain: erba transforming protein homology <ERBA>

Query Match 38.3%; Score 1277; DB 2; Length 563;  
Best Local Similarity 46.6%; Pred. No. 4.7e-63;  
Matches 297; Conservative 70; Mismatches 145; Indels 126; Gaps 20;

QY 40 MDLSTETATATLSLPSISTFVE-GYSSNVELKPCVYQMQRPLIKVEGRAPSYHHHH 98  
DB 1 MDLSPETAPTATPLSFSTFMDGGYTGFEFD---TFLYQLFGTGTQPCSSASST----- 52

QY 99 HHHHHHHHHQHQHQPSPSPASS-----PEDEVLPSTSMYFKQSP 139  
DB 53 -----SSSSSATSPASASPKFEDFQVYGCYPTLSCGLDETSSSGSEYGGSP 100

QY 140 PSTPT--TPAPPPQAGALWD-----EALPSAPGCIAPGLLDPMPK 178  
DB 101 CSAPSPPTNPQPSQLSPWDGSGFHFSPQTYEGLVWTEQLPKASG-----PPP-- 150

QY 179 AVPTVAGARFPLFHPKSPHPAPSPAGGHHLGVDPTAAALSPLGAAAAAGSQAAL 238  
DB 151 -----PPTFFSFSPPTGSPSP-----LAQSSKL---FPAPATHQLGSG 186

QY 239 ESHPYGLPLAKRAAPLAPPLGLT-PSPTASSLGSPLSPSPSSSSSG--EGTCVAVC 295  
DB 187 ESY-----SVPAAPFGLAPTSNCDTSGIL-DAP-VTSKTSRSGSSGSEGCACV 235

QY 296 GDNAAACHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFKCLSV 355  
DB 236 GDNASCQHYGVRTCEGCKGFFKRTVQNAKYVCLANKNCVDRKRRNRQYCRFKCLAV 295

QY 356 GMVKEVVRTDSLKGRGRGLPSKPKPLQEPSPSPSPPICMNALVRALTDSTP--D 413  
DB 296 GMVKEVVRTDSLKGRGRGLPSKPKQP-----PDASPTNLLTSILRAHLDSPGNTAK 346

QY 414 LDYSRYCPTDQAAAAG-TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLES 472  
DB 347 LDYSRYCPTDQAAAAG-TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTDLPKEDQTLLES 406



Qy	473	AFLELFVLR	SIRSWTAEDKVF	FCNGVLVLR	LOCLR	GFGEWLDST	KDPSLNTQSLN	LDIQ	532	
									:	
Db	407	AFLELFILRL	AYRSPKGGKLI	FCSGVLVLR	LQCAR	GFQDWT	DNILATFSRL	SHSLGVDVP	466	
									:	
Qy	533	ALACLSAL	SMITERHGL	KEPKRV	VELCNKNT	TSLLKHDS	KGAALPTE---	SKVLGALV	588	
									:	
Db	467	AFACLSAL	VLITDRHGL	QDPRV	VELQNR	IASCL	KEHNA-AVAGD	QPASCLSR	LGKUP	525
									:	
Qy	589	ELRKICTL	GLQRI	FVYK	LEDLV	SPSII	DKFLDTLPF		626	
									:	
Db	526	ELRLCTGL	QRI	FCLE	LDVPP	PIVDK	IFMDT	LSF	563	
									:	

## RESULT 11

S58238  
DHR38 protein - fruit fly (*Drosophila melanogaster*)  
C/Species: *Drosophila melanogaster*  
C/Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 16-Aug-2004  
C/Accession: S58238  
R/Sutherland, J.; Korlova, T.; Tzertzinis, G.; Kafatos, F.C.  
Submitted to the EMBL Data Library, June 1995  
A/Description: DHR38: a new partner for USP suggests an unexpected role for NGFI-B type  
A/Reference number: S58205  
A/Accession: S58238  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-551 <UT>  
A/Cross-references: EMBL:X89246; NID:G929561; PIDN:CAA61534.1; PID:G929562  
C/Genetics:  
A/Gene: FlyBase:Hz38  
A/Cross-references: FlyBase:FBgn0014859  
C/Superfamily: erba transforming protein homology  
C/Keywords: zinc finger  
F/220-467/Domain: erba transforming protein homology <ERBA>

[illegible]

## RESULT 12

S58205  
DHR38 protein - silkworm (fragment)  
C:Species: Bombyx mori (silkworm)  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 16-Aug-2004  
C:Accession: S58205  
R:Sutherland, J.; Kozlova, T.; Tzertzinis, G.; Kafatos, F.C.  
submitted to the EMBL Data Library, June 1995  
A:Description: DHR38: a new partner for USP suggests an unexpected role for NGPI-B type  
A:Reference number: S58205  
A:Accession: S58205  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-360 <SUT>  
A:Cross-references: UNIPROT:P49870; EMBL:X89247; NID:g929002; PIDN:CAA61535.1; PID:g929029  
C:Superfamily: erBA transforming protein homology  
C:Keywords: zinc finger  
F:24-276/Domain: erBA transforming protein homology <ERBA>

## RESULT 13

RESOLUT 13  
 T43348  
 nuclear steroid hormone receptor NHR-6 - *Caenorhabditis elegans*  
 N;Alternate names: protein I48D5.1; steroid hormone receptor family member CNR8  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 03-Mar-2000 #sequence\_revision 03-Mar-2000 #text\_change 09-Jul-2004  
 C;Accession: T43348; T20039; I45068  
 R;Sluder, A.E.; Mathews, S.W.; Hough, D.; Yin, V.P.; Maina, C.V.  
 Genome Res. 9, 103-120, 1999

A:Title: The nuclear receptor superfamily has undergone extensive proliferation and divergence  
A:Reference number: Z22443; MUID:99148134; PMID:10022975  
A:Accession: T43348  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-619.<SLU>  
A:Cross-references: UNIPROT:P41829; EMBL:AF083224; NID:G4139073; PIDN:AA03682.1; PID:94  
R:Lightning, J.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19215  
A:Accession: T20039  
A:Molecule type: DNA  
A:Residues: 58-619.<WIL>  
A:Cross-references: EMBL:Z36237; PIDN:CAA85271.1; GSPDB:GN00021; CESP:C48D5.1  
A:Experimental source: clone C48D5  
R:Kostrouch, Z.; Kostrouchova, M.; Rall, J. E.  
Proc. Natl. Acad. Sci. U.S.A. 92, 156-159, 1995  
A:Title: Steroid/thyroid hormone receptor genes in *Caenorhabditis elegans*.  
A:Reference number: A55595; MUID:95116514; PMID:7816808  
A:Accession: I45068  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 157-187, 'r', 189-619.<RES>  
A:Cross-references: EMBL:U13076; NID:G538372; PIDN:AAA96984.1; PID:G538373  
C:Genetics:  
A:Gene: nhr-6; cnr8; CESP:C48D5.1  
A:Map position: 3  
A:Introns: 71/1; 98/3; 124/3; 175/2; 240/2; 337/1; 389/3; 510/1; 550/3  
C:Superfamily: *Caenorhabditis elegans* nuclear steroid hormone receptor NHR-6; *erba* trans-  
form protein  
F:266-507/Domain: *erba* transforming protein homology <ERBA>  
  
Query Match 16.6%; Score 554.5; DB 1; Length 619;  
Best Local Similarity 28.4%; Pred. No. 2.4e-23;  
Matches 190; Conservative 79; Mismatches 249; Indels 151; Gaps 24;  
  
QY 4 VQAQVSPSPGSSYAAQTVSSYVTEIMNPVTKLMDLGSTETATATSLPSITFVE 63  
DB 11 LQDQFENCSPAS--VDSSYSSCSVEDEIEIYTRLVN-----EEPLRDRFFR 56  
QY 64 GYSSNVELKPCVY-----QWQPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHH 117  
DB 57 EMSKNSSCSSTFDYGFSGSSSRKSGKTTDADLSL-----FHLVTSQVNTV 107  
QY 118 PRASSPEDEVLPSTSMYFKQSPPTTPPAFPQAGALWDEALPSAGPGCIAP--GFLDLP 175  
DB 108 PRPTKTEVESIEE---FQKPSSS-----SHRLPSEMNASITHIKSELDP 150  
QY 176 PMKA-----VPTVAGARPLF--HFKSPSPHPAPSPAGGHLGYDPTAAALSL 223  
DB 151 TMOAFQMPHNDLFLATAAPHYNPFALSNDFMNPMPSTSPFPQH-----F 198  
QY 224 PLGAAAAGSQAALSHHPYGLPLAKRAAPLAPPLGLTPSPASSLLGESP----- 275  
DB 199 FV-SDRRSGSGTSSN-----NTGGTPSPHSSSLTPSPLOQFLRS 241  
QY 276 -----SLPSP-----PSRSSSGEGTCVAGGNAAQCHYGVRTCEGCKGFFKRTVQKNK 325  
DB 242 FLNPDNLSTPTSGVSEETALDADKCAVNCNDRAVCLHYGARTCEGCKGFFKRTVQKNK 301  
QY 326 YVCLANKNCVDKRRNRCQYCRFOKCLSVGMVKEVVRTDSLKGRLPSPKSPLOQE 385  
DB 302 YTCAGNKTCPIDKRRYRSRCQYCRQKCLEVGMVKEIVRHGSLSGRRGLSSKTK--LARS 359  
QY 386 PSQSPSPSPICMNNALVRALTSTPRDLDSRYCPTDQAAAGTDAEHVQQFVNLITASI 445  
DB 360 EDQSPSPPLLLALMG---KAIEDHTNMTVVRQFMQPFDETI-----LRLHGL 406  
QY 446 DVRSWAEEKIPGTDLPKEDQTLLESFALEFLVRLSIRSNATDAEDKVFVFCNGLVHLRQ 505  
DB 407 HATKLLMAMPQISEIQPADFQILLSRSFPAIMAIRVANRCNSTDTIMFESG-----ELF 462  
QY 506 CLRGFGE-----WLDISKDFSLNQLSLNLDIOALACLALSMT---TERH--GLKEP 552

DB 463 SLNAPPACFOQIRFMVDKARTFS---SLVDWEPQAFAPFALQFLAGNTHVNLGLTNK 519  
QY 553 KRVEELCNKTSLSKDHQSKGQALEPTESTKVGALVELRKICTGLQRIFFYLKLEDLVSP 612  
DB 520 PLVDQVOSTINALKDHCSGSO---NKLAKIVRLTQDFVDFHALGLQAL-----DILYP 570  
QY 613 PSIIDKLFL 621  
DB 571 SHQLPEEFM 579  
  
RESULT 14  
G88408  
protein nhr-6 [imported] - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 17-May-2002  
C:Accession: G88408  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog-  
ical processes  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele-  
gans/  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: G88408  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-562.<STO>  
A:Cross-references: GB:chr\_III; PIDN:CAA85271.1; PID:G3875028; GSPDB:GN00021  
C:Genetics:  
A:Gene: nhr-6  
A:Map position: 3  
C:Superfamily: *Caenorhabditis elegans* nuclear steroid hormone receptor NHR-6; *erba* trans-  
form protein  
  
Query Match 16.6%; Score 554; DB 2; Length 562;  
Best Local Similarity 30.1%; Pred. No. 2.3e-23;  
Matches 169; Conservative 72; Mismatches 199; Indels 122; Gaps 20;  
  
QY 105 HHHHQQHQHQPSIPASSPEDEVLPSTSMYFKQSPPTTPPAFPQAGALWDEALPSAP 164  
DB 38 HSLVTSQVNTVPKPTKTEVESIEE---FEQKPSSS-----SHRLPSEM 80  
QY 165 GCIAIP--GFLDLPMPKA-----VPTVAGARPLF--HFKSPSPHPAPSPAGGHH 210  
DB 81 NASITHIKSELDPMTQAFQMPHNDLFLATAAPHYNPFALSNDFMNPMPSTSPFPQH 140  
QY 211 LGYDPTAAALSLPLGAAAAGSQAALSHHPYGLPLAKRAAPLAPPLGLTPSPASSLL 270  
DB 141 -----FV-SDRRSGSGTSSN-----NTGGTPSPHSSSL 171  
QY 271 LGESP-----SLPSP-----PSRSSSGEGTCVAGGNAAQCHYGVRTCEGC 312  
DB 172 PTSPPLOQFLRSFLNPDNLSTPTSGVSEETALDADKCAVNCNDRAVCLHYGARTCEGC 231  
QY 313 KGFRTVQKNKCYVCLANKNCVDKRRNRCQYCRFOKCLSVGMVKEVVRTDSLKGRRG 372  
DB 232 KGFRTVQKNKCYVCLANKNCVDKRRNRCQYCRFOKCLSVGMVKEVVRTDSLKGRRG 291  
QY 373 RLSPKPSPLQOEPSPSPSPPICMNNALVRALTSTPRDLDSRYCPTDQAAAGTDAE 432  
DB 292 RLSSKTK--LARSEDPSPPLLLALMG---KAIEDHTNMTVVRQFMQPFDETI----- 341  
QY 433 HVQPFYNLLTASIDVRSWAEEKIPGTDLPKEDQTLLESFALEFLVRLSIRSNATDAEDK 492  
DB 342 -----LRLHGLHATKLLMAMPQISEIQPADFQILLSRSFPAIMAIRVANRCNSTDT 396  
QY 493 FVFCNGLVHLRQCLRGFGE-----WLDISKDFSLNQLSLNLDIOALACLALSMT- 543  
DB 397 IMFESG-----ELFSLNAPFACFOQIRFMVDKARTFS---SLVDWEPQAFAPFALQFLA 449  
QY 544 --TERH--GLKEPFRVEELCNKTSLSKDHQSKGQALEPTESTKVGALVELRKICTGLQ 599  
DB 450 GNTHEVNLGLTNKPLVDQVOSTINALKDHCSGSO---NKLAKIVRLTQDFVDFHALGLQ 506

Job time : 44 secs

QY 600 RIFYLKLELVSPPSIIDKLEL 621  
Db 507 AL-----DILYPSHQLPEPFM 522

RESULT 15  
S13512  
retinoic acid receptor beta-1 - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 21-Nov-1993 #sequence\_revision 22-Apr-1995 #text\_change 09-Jul-2004  
C:Accession: S13512; S17073  
R:Padanilam, B.J.; Mcleod, L.B.; Suzuki, H.; Solursh, M.  
Nucleic Acids Res. 19, 395, 1991  
A:Title: Nucleotide sequence of an isoform of chicken retinoic acid binding protein-beta  
A:Reference number: S13512; MUID:91195065; PMID:1849630  
A:Accession: S13512  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-455 <PAD>  
A:Cross-references: UNIPROT:P22448; EMBL:X56674; NID:G63785; PIDN:CAA39997.1; PID:G63786  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1990  
R:Nohno, T.; Muto, K.; Noji, S.; Saito, T.; Taniguchi, S.  
Biochim. Biophys. Acta 1089, 273-275, 1991  
A:Title: Isoforms of retinoic acid receptor beta expressed in the chicken embryo.  
A:Reference number: S16243; MUID:91274365; PMID:1647216  
A:Accession: S17073  
A:Molecule type: mRNA  
A:Residues: 1-29, 'P', 31-80 <NOH>  
C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology  
C:Keywords: alternative splicing; DNA binding; nucleus; transcription regulation; zinc f  
F:86-337/Domain: erba transforming protein homology <ERBA>  
F:88-108/Region: zinc finger  
F:124-148/Region: zinc finger

Query Match 16.3%; Score 543; DB 2; Length 455;  
Best Local Similarity 32.5%; Pred. No. 7 4e-23;  
Matches 149; Conservative 65; Mismatches 171; Indels 74; Gaps 14;

QY 180 VPTVAGARPLPHFKPSPHPAPSPAGGHILGYDPTAAALSLPLGAAAAAGSQAALALE 239  
Db 10 VPAVNG---HMTHY-PAAPYPLLFAPIVIG-----GLSLP-----SLHGLQ 45

QY 240 SHPYGLPLAKRAAPLAPPLGL-TPSP-----TASSLLGESPSLSPSPSSSSSGEGT 291  
Db 46 SHP-----PTSGCSTPSATVETQSTSEELVSPSPSPPLPPRVYK----P 87

QY 292 CAVCGDNAACQHVGRTCGCGFFKRTYQKNKAKVCLANKKNCVDRKRNRNCOYCRFOK 351  
Db 88 CFVCQDKSSGYHYGVSAACGCGKGFRRSIQKNMVYTCRDKNCVINKVTRNRCQYCRLOK 147

QY 352 CLSVGMVKEVVRVTDLSLKGRRRLPSKPGSPLOQEPSQSPSPSPPICMNNALVRALTDSTP 411  
Db 148 CFEVGMKESVNRNKK-----KKKEPTQESTENYEMTAELDDLTEKIRKAHQET- 198

QY 412 RDLVSVRYCPTQAAAGTDAEH-----VOQFYNLLTASIDVSRSWAEKIPGFTDLPK 463  
Db 199 ----FPSLCQLGKYTTNSADHRVRLDLGLWDKFSGLATKCIKIVEFAKRLPGFTSLTI 254

QY 464 EDQTLIESAFLEFLVRLSIRNTAEDKFCVNCGLVLRLOCLR-GFGEWLDSIKDFSL 522  
Db 255 ADQITLLKAACLDILRLICTRYTPEDTWTFSQDTNLTNRQTMHNAFGFLDVLVFTFAN 314

QY 523 NQLSNLMDTQALACLSALSMIT-ERHGLKEPKRVEELCNKITSSLNKHOSKGOALEP-TE 580  
Db 315 QLLPLEMDDETGLLSAICLICDRQDRLDEPKVKYDQLQEPLEALKIYIRKRPKNKPHMF 374

QY 581 SKVLGALVELRKICTLGLQRIFYFLKLELVSPPSIIDKL 619  
Db 375 PKILMKITDLRSISAKGAERVITLKWEIFGSMPLIQEM 413

Search completed: March 12, 2005, 17:52:00

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 17:40:16 ; Search time 176 Seconds  
(without alignments)  
1821.373 Million cell updates/sec

Title: US-10-608-863-2

Perfect score: 3337

Sequence: 1 MPCVQAQYSPSPGSSYAAQ.....EDLVSPPSIIDKFLDLPF 626

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	3333	99.9	626	1 NR43 HUMAN	Q92570 homo sapien
2	3133.5	93.9	643	2 O97726	Q97726 sus scrofa
3	3074	92.1	628	1 NR43 RAT	P51179 rattus norv
4	3054.5	91.5	627	1 NR43 MOUSE	Q9qzb6 mus musculu
5	2099.5	62.9	446	2 O97727	Q97727 sus scrofa
6	1751	52.5	598	1 NR42 HUMAN	P43354 homo sapien
7	1746	52.3	598	1 NR42 MOUSE	Q06219 mus musculu
8	1736	52.0	598	1 NR42 RAT	Q07917 rattus norv
9	1677	50.3	600	2 Q98TQ3	Q98TQ3 oryzias lat
10	1598	47.6	586	2 Q6DH08	Q6dh08 brachydanio
11	1572	47.1	535	2 Q6NXU0	Q6nxu0 homo sapien
12	1371	41.1	598	1 NR41 CANFA	P51666 canis famli
13	1368	41.0	598	2 Q8N3V2	Q8n3v2 homo sapien
14	1367	41.0	598	1 NR41 HUMAN	P22736 homo sapien
15	1364	40.9	598	2 Q61B08	Q61bu8 homo sapien
16	1353	40.5	652	2 Q6ZMM6	Q6zmm6 homo sapien
17	1351.5	40.5	577	2 Q61NY4	Q61ny4 xenopus lae
18	1350.5	40.5	577	1 NR42 XENLA	Q04913 xenopus lae
19	1350.5	40.5	577	2 Q77T0V3	Q77t0v3 xenopus lae
20	1322.5	39.6	597	1 NR41 RAT	P22829 rattus norv
21	1316.5	39.5	601	1 NR41 MOUSE	P12813 mus musculu
22	1254	37.6	574	2 Q6GMG3	Q6gmg3 brachydanio
23	1219.5	36.5	808	2 Q704C6	Q7q4c6 anopheles g
24	1197.5	35.9	545	2 Q9U4L1	Q9u4l1 aedes aegypt
25	1190.5	35.7	1073	1 HR38 DROME	P49869 drosophila
26	1168.5	35.0	556	2 Q81NU7	Q81nu7 drosophila
27	1057.5	31.7	392	2 Q9R1W4	Q9r1w4 mus musculu
28	1023	30.7	360	1 HR38 BOMMO	P49870 bombyx mori
29	554.5	16.6	619	1 NR46 CAEEL	P41829 caenorhabdi
30	545	16.3	455	1 RRB_CHICK	P22448 gallus gall
31	545	16.3	455	1 RRB_COTJA	Q9w6b3 coturnix co

32	541	16.2	156	2 Q6AW80	Q6aw80 homo sapien
33	541	16.2	482	1 RRB_MOUSE	P22605 mus musculu
34	533.5	16.0	533	1 RXRB_HUMAN	P28702 homo sapien
35	531.5	15.9	462	1 RRA_MOUSE	P11416 mus musculu
36	531.5	15.9	462	2 P97513	P97513 mus spretus
37	531	15.9	455	1 RRB_HUMAN	P10826 homo sapien
38	530.5	15.9	462	1 RRB_HUMAN	P10276 homo sapien
39	529.5	15.9	525	2 Q95L53	Q95l53 mustela vis
40	527	15.8	590	2 Q6P2H5	Q6p2h5 homo sapien
41	527	15.8	600	2 Q6IN48	Q6in48 homo sapien
42	523.5	15.7	460	1 RRA_CHICK	Q90966 gallus gall
43	521.5	15.6	520	1 RXRB_MOUSE	P28704 mus musculu
44	520.5	15.6	458	1 RRA_NOTVI	P18514 notophthalm
45	519.5	15.6	451	2 Q6MGB3	Q6mgb3 rattus norv

#### ALIGNMENTS

#### RESULT 1

ID NR43 HUMAN STANDARD; PRT; 626 AA.  
AC Q92570; Q12935; Q14979; Q16420; Q9UEK2; Q9UEK3;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Orphan nuclear receptor NR4A3 (Nuclear hormone receptor NOR-1)  
DE (Neuron-derived orphan receptor 1) (Mitogen induced nuclear orphan receptor).  
GN Name=NR4A3; Synonyms=CHN, CSMF, MINOR, NOR1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RC TISSUE=Fetal brain;  
RX MEDLINE=96404972; PubMed=8809112; DOI=10.1016/0167-4781(96)00101-7;  
RA Ohkura N., Ito M., Tsukada T., Sasaki K., Yamaguchi K., Miki K.;  
RT "Structure, mapping and expression of a human NOR-1 gene, the third member of the Nur77/NGFI-B family.";  
RL Biochim. Biophys. Acta 1308:205-214(1996).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RC TISSUE=Periphereal blood;  
RX MEDLINE=96192925; PubMed=8614405; DOI=10.1210/me.9.12.1692;  
RA Hedvat C.V., Irving S.G.;  
RT "The isolation and characterization of MINOR, a novel mitogen-inducible nuclear orphan receptor.";  
RL Mol. Endocrinol. 9:1692-1700(1995).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RC TISSUE=Fetal brain;  
RX MEDLINE=96152889; PubMed=8570200;  
RA Clark J., Benjamin H., Gill S., Sidhar S., Goodwin G., Crew J., Gusterson B.A., Shipley J., Cooper C.S.;  
RT "Fusion of the EWS gene to CHN, a member of the steroid/thyroid receptor gene superfamily, in a human myxoid chondrosarcoma.";  
RL Oncogene 12:229-235(1996).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM BETA), AND CHROMOSOMAL TRANSLOCATION WITH EWS.  
RC TISSUE=Fetal heart;  
RX MEDLINE=96177652; PubMed=8634690;  
RA Labelle Y., Zucman J., Steman G., Kindblom L.-G., Knight J., Turc-Carel C., Dockhorn-Dworniczak B., Mandahl N., Desmaza C., Peter M., Aurias A., Delattre O., Thomas G.;  
RT "Oncogenic conversion of a novel orphan nuclear receptor by chromosome translocation.";  
RL Hum. Mol. Genet. 4:2219-2226(1995).  
RN [5]  
RP SEQUENCE OF 1-69 AND 301-443 FROM N.A. (ISOFORM BETA), AND ALTERNATIVE SPLICING.

RC TISSUE=Skeletal muscle;  
 RX MEDLINE=98241504; PubMed=9573341; DOI=10.1016/S0378-1119(98)00095-X;  
 RA Ohkura N., Ito M., Tsukada T., Sasaki K., Yamaguchi K., Miki K.;  
 RT "Alternative splicing generates isoforms of human neuron-derived  
 orphan receptor-1 (NOR-1) mRNA.";  
 RL Gene 211:79-85(1998).  
 CC -!- FUNCTION: Binds to the BIA response-element.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Alpha;  
 CC IsoId=Q92570-1; Sequence=Displayed;  
 CC Name=Beta;  
 CC IsoId=Q92570-2; Sequence=VSP\_003712; VSP\_003713;  
 CC TISSUE SPECIFICITY: High expression of isoform alpha in skeletal  
 CC muscle. High expression of isoform beta in skeletal muscle and low  
 CC expression in fetal brain and placenta.  
 CC -!- DISEASE: Involved in Ewing's sarcoma through chromosomal  
 CC translocations t(9;22)(q22-31;q11-12) which involve EWS and NR4A3.  
 CC -!- DISEASE: Involved in a form of extra-skeletal myxoid  
 CC chondrosarcomas (EMC) through a chromosomal translocation  
 CC t(9;17)(q22;q11) that involves TAF2N and NR4A3  
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4  
 CC subfamily.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts  
 CC in positions 4 and 20.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D78579; BAAL1419.1; -;  
 CC EMBL; U12767; AAB02581.1; ALT\_FRAME.  
 CC EMBL; S81243; AAB36006.1; ALT\_INIT.  
 CC EMBL; X89894; CAA61984.1; -;  
 CC EMBL; D85241; BAA28608.1; -;  
 CC EMBL; D85242; BAA31221.1; -;  
 CC HSP; P43354; IOVL.  
 CC TRANSFAC; T04750; -;  
 CC TRANSFAC; T04753; -;  
 CC Genew; HGNC:7982; NR4A3.  
 CC MIM; 600542; -;  
 CC GO; GO:0005488; F.binding; TAS.  
 CC GO; GO:0003707; F.steroid hormone receptor activity; TAS.  
 CC GO; GO:0004887; F.thyroid hormone receptor activity; TAS.  
 CC InterPro; IPR00536; Hrmn\_recept\_lig.  
 CC InterPro; IPR001723; Stdhrn\_receptor.  
 CC InterPro; IPR008946; Str\_ncl\_receptor.  
 CC InterPro; IPR001628; Znf\_C4steroid.  
 CC Pfam; PF00104; Hormone\_recep; 1.  
 CC Pfam; PF00105; zf-C4; 1.  
 CC PRINTS; PR00398; STRDHORMONER.  
 CC PRINTS; PR00047; STROIDFINGER.  
 CC ProDom; PD000035; Znf\_C4steroid; 1.  
 CC PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 CC Alternative splicing; Chromosomal translocation; DNA-binding;  
 CC Nuclear protein; Proto-oncogene; Receptor; Transcription regulation;  
 CC Zinc-finger.  
 CC FT DNA\_BIND 292 357 Nuclear receptor-type.  
 CC FT ZN\_FING 292 312 C4-type.  
 CC FT ZN\_FING 328 352 C4-type.  
 CC FT DOMAIN 440 490 Ligand-binding (Potential).  
 CC FT DOMAIN 95 108 Poly-His.  
 CC FT DOMAIN 282 287 Poly-Ser.  
 CC FT VARSPLIC 419 443 YCPTQAAAGTDAEHVQOQFYNLLTA -> VSFMISCFQWMD  
 CC QGLYLWLLVIRYD (in isoform Beta).  
 CC /FTId=VSP\_003712.  
 CC Missing (in isoform Beta).  
 CC /FTId=VSP\_003713.  
 CC FT VARSPLIC 444 626 .

FT CONFLICT 240 240 G -> S (in Ref. 3 and 4).  
 FT CONFLICT 454 454 K -> R (in Ref. 1).  
 FT CONFLICT 579 579 T -> N (in Ref. 1).  
 FT CONFLICT 585 585 G -> V (in Ref. 1).  
 SQ SEQUENCE 626 AA; 68199 MW; 40D18DA8FEF991EC CRC64;  
 Query Match 99.9%; Score 3333; DB 1; Length 626;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-168;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MPCVQAQYSPSPGSSYAAQYSSYITTEINPDYTKLTMDLGSTEITATATSLPSIST 60  
 DB 1 MPCVQAQYSPSPGSSYAAQYSSYITTEINPDYTKLTMDLGSTEITATATSLPSIST 60  
 QY 61 FVEGVSSNVELKPCVCYQMQRPDIKVERGRAPSYHHHHHHHHHHHHHHHHHHHHH 120  
 DB 61 FVEGVSSNVELKPCVCYQMQRPDIKVERGRAPSYHHHHHHHHHHHHHHHHHHHHH 120  
 QY 121 SSPDEVLPSTSMYFKQSPSPSTPTTTPAPPPQAGALWDEALPSAPGCIAPGPIIDPPMKAV 180  
 DB 121 SSPDEVLPSTSMYFKQSPSPSTPTTTPAPPPQAGALWDEALPSAPGCIAPGPIIDPPMKAV 180  
 QY 181 PTVAGARPLFHFKPSPHPHPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALLES 240  
 DB 181 PTVAGARPLFHFKPSPHPHPAPSPAGGHHLGYDPTAAALSLPLGAAAAAGSQAALLES 240  
 QY 241 HPYGLPLAKRAAPLAFPLGLTPSPASSLLGESPLSPSPSSSSGEGTCVCGDNAA 300  
 DB 241 HPYGLPLAKRAAPLAFPLGLTPSPASSLLGESPLSPSPSSSSGEGTCVCGDNAA 300  
 QY 301 COHGVRTCEGKGFFKRTVOKNAYVCLANKNCVPDKRRNRNRCQYCRFKCLSVGMVKE 360  
 DB 301 COHGVRTCEGKGFFKRTVOKNAYVCLANKNCVPDKRRNRNRCQYCRFKCLSVGMVKE 360  
 QY 361 VVRTSLKGRGRPLSKPKSLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420  
 DB 361 VVRTSLKGRGRPLSKPKSLQBPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 420  
 QY 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480  
 DB 421 PTDQAAAGTDAEHVQOQFYNLLTASIDVSRSWAEKIPGFTDLPKEDQTLIIESAFLEFVL 480  
 QY 481 RLRSRNTAEDKVFVFCNGLVLRHQLQCLRGFGWELDSIKDFSLNLSLNLDIQALCLSA 540  
 DB 481 RLRSRNTAEDKVFVFCNGLVLRHQLQCLRGFGWELDSIKDFSLNLSLNLDIQALCLSA 540  
 QY 541 SMITRHHGLKPKRVEELCNKITSSLDKDHQSGQALEPTESKVLGALVELRKICTIGLQR 600  
 DB 541 SMITRHHGLKPKRVEELCNKITSSLDKDHQSGQALEPTESKVLGALVELRKICTIGLQR 600  
 QY 601 IFYKLEDLVSPPSIIDKLFDTLPF 626  
 DB 601 IFYKLEDLVSPPSIIDKLFDTLPF 626  
 RESULT 2  
 O97726 PRELIMINARY; PRT; 643 AA.  
 ID O97726  
 AC O97726;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Neuron-derived orphan receptor-1 alfa.  
 GN Name=NOR-1;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cases-Langhoff C., Castello A., Martinez-Gonzalez J., Badimon L.;  
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).

```

CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL: AJ011767; CA009763.1; -.
DR HSSP: P43354; LOVL.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR000536; Hrmn_recept_lig.
DR InterPro: IPR003072; NucR_receptor.
DR InterPro: IPR003070; NOC_receptor.
DR InterPro: IPR001723; Stdhrmn_receptor.
DR InterPro: IPR008946; Str_ncl_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; Hormone_recep; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR01286; NORNURCEPTR.
DR PRINTS: PR01284; NUCLEARRECPTR.
DR PRINTS: PR00398; STRODHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 643 AA; 70077 MW; 20596A4222FB4CA7 CRC64;

Query Match 93.9%; Score 3133.5; DB 2; Length 643;
Best Local Similarity 92.4%; Pred. No. 5e-158;
Matches 594; Conservative 10; Mismatches 22; Indels 17; Gaps 3;

QY 1 MPCVQAQYSPSPGSSYAAQTYSSEYTTIMPDYTKLTMDLGSITATATSLPSIST 60
DB 1 MPCVQAQYSPSPGSSYAAQTYSSEYTTIMPDYTKLTMDLGSITATATSLPSIST 60
QY 61 FVEGVSSNVELKPCVYQMO---RPLIKVEGRAPSHHHHHHHHHHHHHHHQO----- 110
DB 61 FVEGVSSNVELKPCVYQMO---RPLIKVEGRAPSHHHHHHHHHHHHHHHHHHHHHQO 120
QY 111 --HQQPSIPASSPEDEVLPTSMYFKQSPSTPTTAPPPQAGALWDEALPSAPGCTA 168
DB 121 PFQOQPSIPSPGPEDEVLPTSMYFKQSPSTPTTTPVPFQAGALWEDALPSAGGCTA 180
QY 169 PGFLDPPMKAVTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLG-- 226
DB 181 PGFLDPPMKAVTVAGAPFLFHFKPSPPHPAPSPAGGHHLYGDPYTAALSLPLGAA 240
QY 227 ---AAAAGSQAALSHPYGLPLAKRAAPLAPPPLGLTPSPASSLLGESPLSPSPSR 283
DB 241 AAAAAAGSQAALSHPYGLPLAKRAAPLAPPPLGLTPSPASSLLGESPLSPSPSR 300
QY 284 SSSSGEGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNR 343
DB 301 STASGEGTCVACGDNAACQHYGVRTCEGCKGFFKRTVQKNKYVCLANKNCVDRKRRNR 360
QY 344 QYCRPQKCLSGVMKVEVTRTDSLKGRRGLRSLKPSPLQOQSPSPSPSPSPSPSPSP 403
DB 361 QYCRPQKCLSGVMKVEVTRTDSLKGRRGLRSLKPSPLQOQSPSPSPSPSPSPSPSP 420
QY 404 RALTDTSTPRDLDSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 463
DB 421 RALTDTSTPRDLDSRYCPTDQAAGTDAEHVQOYFNLLTASIDVSSSWAEKIPGFTDLPK 480
QY 464 EQOTLLIESAFLELVRLSIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLSDSIKDFSLN 523
DB 481 EQOTLLIESAFLELVRLSIRNTAEDKVFVFCNGLVLRHLOCLRGFGWLSDSIKDFSLR 540
QY 524 LQSLNLDIQAALCLSALSALSMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTSKV 583
DB 541 LQSLNLDIQAALCLSALSALSMITERHGLKEPKRVEELCNKITSSLKDHQSKQALEPTPKV 600
QY 584 LGALVELRKICTGLQRIYFKLEDLVSPSPSIIDKFLDLP 626
DB 584 LGALVELRKICTGLQRIYFKLEDLVSPSPSIIDKFLDLP 626

```





```

FT  SQ  SEQUENCE  627 AA;  68454 MW;  2F49CB4C8DE315A8 CRC64;  /FTIG=VSP 010084.
Query Match  91.5%;  Score 3054.5;  DB 1;  Length 627;
Best Local Similarity  91.8%;  Pred. No. 7.4e-154;
Matches 579;  Conservative 12;  Mismatches 31;  Indels 9;  Gaps 4;

QY  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTMMNPDYTKLTMDLGSSTETATATSLPSIST 60
DB  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTMMNPDYTKLTMDLGSSTETATATSLPSIST 60

QY  61  FVEGYSSNYELKPSVCYQWQ----QRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 116
DB  61  FVEGYSSNYELKPSVCYQWQ----QRPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 117

QY  117  IPPASSPEDEVLPSTSMYFKQSPSTPTTTPAFPPQAGALWDEALPSAPGCIAPGILLDP 176
DB  118  IPPASSPEDEVLPSTSMYFKQSPSTPTTTPAFPPQAGALWDEALPSAPGCIAPGILLDP 177

QY  177  MKAVPTV-AGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPYTAALSLPLGAAAAGSQA 235
DB  178  MKAVPMPAAAAPFPIF-FKPSPPHPPAPSPAGGHHLYGDPYTAALSLPLGAAAAGSQA 236

QY  236  AALESHPYGLPLAKRAAPLAFPLGLTPTASSLLGESPLSPSPSSSGEGTCAVC 295
DB  237  AALEGHPYGLPLAKRTATLTFFPLGLTASPTASSLLGESPLSPSPNRSSSSGEGTCAVC 296

QY  296  GNAACQHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNRQCYCRFQKCLSV 355
DB  297  GNAACQHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNRQCYCRFQKCLSV 356

QY  356  GMVKEVVRTSLKGRRLPSKPSPLQOEPPSPSPPICMNALVRALTDSTPRDL 415
DB  357  GMVKEVVRTSLKGRRLPSKPSPLQOEPPSPSPPICMNALVRALTDSTPRDL 416

QY  416  YSRYPCTDQAAGTDAEHVQYQFNLLTASIDVSRSWAEKI PGFTDLPKEDQTLLESAPL 475
DB  417  YSRYPCTDQAAGTDAEHVQYQFNLLTASIDVSRSWAEKI PGFTDLPKEDQTLLESAPL 476

QY  476  ELFVLRLSIRNSTAEDKVFVNCGLVLRLQCLRGFGEWLDSIKDFSLNLQSLNLDIQALA 535
DB  477  ELFVLRLSIRNSTAEDKVFVNCGLVLRLQCLRGFGEWLDSIKDFSLNLQSLNLDIQALA 536

QY  536  CLSALSMITERGLKPEKREBELCNKITSSLDKHQSKGOALEPTESKVLGALVELRKICT 595
DB  537  CLSALSMITERGLKPEKREBELCNKITSSLDKHQSKGOALEPTESKVLGALVELRKICT 596

QY  596  LGLQRFYKLKEDLVSPPSIIBKLFDTLUPF 626
DB  597  LGLQRFYKLKEDLVSPPSIIBKLFDTLUPF 627

RESULT 5
O97727  PRELIMINARY;  PRT;  446 AA.
AC  O97727;
DT  01-MAY-1999 (TrEMBLrel. 10, Created)
DT  01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Neuron-derived orphan receptor-1 beta.
GN  Name=NOR-1;
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
RN  [1]_
SEQUENCE FROM N.A.
RA  Cases-Ganghoff C.; Castello A.; Martinez-Gonzalez J.; Badimon L.;
RL  Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC  -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC  -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR  EMBL; AJ011768; CAA09764.1; -.
DR  HSSP; P43354; 10VL.

```

---

```

DR  GO; GO:0005634; C:nucleus; IEA.
DR  GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
DR  GO; GO:0003700; F:transcription factor activity; IEA.
DR  GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR  InterPro; IPR003072; NOR1_receptor.
DR  InterPro; IPR003070; Nucoriph_receptor.
DR  InterPro; IPR001628; Znf_C4steroid.
DR  Pfam; PF00105; Zf-C4; 1_C4steroid.
DR  PRINTS; PR01286; NORNUCARECTR.
DR  PRINTS; PR01284; NUCLEARECTR.
DR  PRINTS; PR00047; STROIDFINGER.
DR  ProDom; PD000035; Znf_C4steroid. 1.
DR  SMART; SM00399; Znf_C4; 1.
DR  PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW  DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW  Transcription regulation; Zinc; Zinc-finger.
SQ  SEQUENCE  446 AA;  47716 MW;  E21960657606FB92 CRC64;

Query Match  62.9%;  Score 2099.5;  DB 2;  Length 446;
Best Local Similarity  89.9%;  Pred. No. 1.8e-103;
Matches 391;  Conservative 9;  Mismatches 18;  Indels 17;  Gaps 3;

QY  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTMMNPDYTKLTMDLGSSTETATATSLPSIST 60
DB  1  MPCVQAQYSPSPGSSYAAQTYSSYTTTMMNPDYTKLTMDLGSSTETATATSLPSIST 60

QY  61  FVEGYSSNYELKPSVCYQWQ----RPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 110
DB  61  FVEGYSSNYELKPSVCYQWQ----RPLIKVEGRAPSYHHHHHHHHHHHHHHHHHHHH 120

QY  111  --OHOQPSIPASSPEDEVLPSTSMYFKQSPSTPTTTPAFPPQAGALWDEALPSAPGCTA 168
DB  121  PPQOQPSIPASSPEDEVLPSTSMYFKQSPSTPTTTPAFPPQAGALWDEALPSAPGCTA 180

QY  169  PGFLDLPDMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPYTAALSLPLGAA 226
DB  181  PGFLDLPDMKAVPTVAGARFPLFHFKPSPPHPPAPSPAGGHHLYGDPYTAALSLPLGAA 240

QY  227  ---AAAAAGSQAALLESHPYGLPLAKRAAPLAFPLGLTPTASSLLGESPLSPSPSR 283
DB  241  AAAAAAGSQAALLESHPYGLPLAKRAAPLAFPLGLTPTASSLLGESPLSPSPSR 300

QY  284  SSSSGEGTCVCGDNACQHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNR 343
DB  301  STASGEGTCVCGDNACQHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDPKRRNR 360

QY  344  COYCRFQKCLSVGMVKEVVVRTDSLKGRRLPSKPSPLQOEPPSPSPPICMNALV 403
DB  361  COYCRFQKCLSVGMVKEVVVRTDSLKGRRLPSKPSPLQOEPPSPSPPICMNALV 420

QY  404  RALTDSTPRDLDSYR 418
DB  421  RALTDSTPRDLDSYR 435

RESULT 6
NR42_HUMAN
ID  NR42_HUMAN  STANDARD;  PRT;  598 AA.
AC  P43354; O16311;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1)
DE  (Immediate-early response protein NOT) (Transcriptionally inducible
DE  nuclear receptor).
GN  Name=NR4A2; Synonyms=NOT, NURR1, TINUR;
OS  Homo sapiens (Human)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN  [1]_
SEQUENCE FROM N.A.
RP  TISSUE=Blood;

```



```

QY 444 SIDVSRWAEKIPGFTDLPKEDQTLIESAFLEFVLRSLRSNTAEDKVFPCNGVLVLR 503
DB 413 SMEIIRGWAELPGFADLPKADODLLFESAFLEFVLRSLRSNPVEGKLIFCNGVVLHR 472
QY 504 LCLRGFGWLSIDKFSNLNLSLNDIOALACLALSALSMITERHGLKEPKRVEELCNKIT 563
DB 473 LQCVRGFGWISDIVEFSSNLQNMIDISAFSCIAALAMVTERHGLKEPKRVEELQNKIV 532
QY 564 SSKLDHQ--SKQOALEPTE-SKVLGALVELRKICTGLQRIPLYLKLEDLVSPSIDKLF 620
DB 533 NCLCKHVTFNNGLNRPNYLSKLLGKLPELRLCTQGLQRIPLYLKLEDLVPPPAIDKLF 592
QY 621 LDTLPF 626
DB 593 LDTLPF 598

RESULT 7
NR42_MOUSE
ID NR42_MOUSE STANDARD; PRT; 598 AA.
AC Q06219; O08690;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1) (NUR-
related factor 1).
GN Name=NR4A2; Synonyms=Nurr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93149122; PubMed=1491694; DOI=10.1210/me.6.12.2129;
RA Law S.W.; Connely O.M.; Denayo F.J.; O'Malley B.W.;
RT "Identification of a new brain-specific transcription factor, NURR1.";
RL Mol. Endocrinol. 6:2129-2135(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=97288523; PubMed=9143501; DOI=10.1006/geno.1997.4677;
RA Castillo S.O.; Xiao Q.; Lyu M.S.; Kozak C.A.; Nikodem V.M.;
RT "Organization, sequence, chromosomal localization, and promoter
identification of the mouse orphan nuclear receptor Nurrl gene.";
RL Genomics 41:250-257(1997).
CC -!- FUNCTION: Probable nuclear receptor that regulates gene expression
in brain tissue.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=NURR1;
CC IsoId=Q06219-1; Sequence=Displayed;
CC Name=NURR1A;
CC IsoId=Q06219-2; Sequence=VSP_003710, VSP_003711;
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- DEVELOPMENTAL STAGE: Expression begins in the embryo, increases in
neonates and decreases in the adult.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR4
subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; S53744; AAB251138.1; -
CC EMBL; U86783; AAC53153.1; -
CC EMBL; U86783; AAC53154.1; -

```

```

DR PIR; A46225; A46225.
DR HSSP; P43354; 10VL.
DR TRANSPAC; T04312; -.
DR MGD; MGI:1352456; Nr4a2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0007399; P:neurogenesis; IMP.
DR GO; GO:0042053; P:regulation of dopamine metabolism; IDA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR01723; Stchrnm_receptor.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STROHOMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Alternative splicing; DNA-binding; Nuclear protein; Receptor;
KW Transcription regulation; Zinc-finger.
FT DOMAIN 80 138
FT PRO-RICH.
FT DNA_BIND 127 233
FT DNA_BIND 263 328
FT ZN_FING 263 283
FT ZN_FING 299 323
FT ZN_FING 346 391
FT DOMAIN 346 391
FT DOMAIN 409 459
FT VARSPLIC 455 455
FT VARSPLIC 456 598
FT Missing (in isoform NURR1A).
FT /FTId=VSP_003710.
FT /FTId=VSP_003711.
SQ SEQUENCE 598 AA; 66592 MW; 66592 MW; DIACSEB2FFB8D366 CRC64;

Query Match 52.3%; Score 1746; DB 1; Length 598;
Best Local Similarity 54.4%; Pred. No. 1.2e-84;
Matches 362; Conservative 75; Mismatches 121; Indels 108; Gaps 18;

QY 1 MPCVQAOVSPSPGGSSYAAQTY-----SSSYTEINMPDYTKTMDLGSTEITATATSLP 56
DB 1 MPCVQAOVSPSPGGSSYAAQTY-----SSSYTEINMPDYTKTMDLGSTEITATATSLP 58
QY 57 STSTFVEGYSSNYELKPSCVVQM-----ORPLIKVEGRAPSYHHHHHHHHHHHHHHHQQQH 112
DB 59 SPSTFMDNYSYGYDVKPCLYQMPLSGQSSIKVEDIQHNTYQQHSH----- 105
QY 113 QQPSIIPPASSPEDEVLP-STSMYFKQSPSPSTTPAFPQQAGALWDEALPSAPGCIAPGP 171
DB 106 ---LPPQS---EEMPHSGSVYKPSPPPTSTPSTFQVQHSPPMDD-----PGS 148
QY 172 LLDPPMKAVPT-----VAGARPLFHFKPSPPHPPA-----PSPAG 207
DB 149 LHNPHQNYVATHTMIETQRTKTPVSRSLFSFKQSPGCTPVSSCCMRFDGFLHVPMPPEPAG 208
QY 208 GHHLGYDPTAAALSLPLGAAAAGSQAALSHSPYGLP-LAKRAAPLAFPLPLGLTPSPPT 266
DB 209 SHHV-----VDGQTFAPVNPTRKPSMGFGQLQI---GH 239
QY 267 ASSLLGESPSLPSPSPSSSGEGTCAVCGDNAAACQHYGVRTCEGCKGFKRTVQKNKY 326
DB 240 ASQLL--DTQVPSPPSRGSPSNEGICAVCGDNAAACQHYGVRTCEGCKGFKRTVQKNKY 297
QY 327 VCLANKNCVDRKRRNRQYCRQKCLSVGMVKEVVRTDLSLKGRRGLRPSKPSKPSLQOQBP 386
DB 298 VCLANKNCVDRKRRNRQYCRQKCLAVGMVKEVVRTDLSLKGRRGLRPSKPSKPSKPSKPSKPS 352
QY 387 SOPSPPSPICMNNALVRALTDSTP--RDLDYSRY-CPTDQAAAGTDAASHVOQFYNLLTA 443
DB 353 QDPSPSPSPVSLISALVRAHVDNSNPAMTSLDYSRFQANPDYQMSGDDTQHIQQOYDILLTG 412
QY 444 STDVSRWAEKIPGFTDLPKEDQTLIESAFLEFVLRSLRSNTAEDKVFPCNGVLVLR 503
DB 413 SMEIIRGWAELPGFADLPKADODLLFESAFLEFVLRSLRSNPVEGKLIFCNGVVLHR 472

```











```

Db      217 QUGER-----SISTAPGL-----APTSPHLDGPGMLDAP-VPSAKARSG 258
QY      286 --SSGEGTCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNR 343
Db      259 APGSGEACVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNR 318
QY      344 CQCFRQKCLSGVMKVEVVRTDSLKGRRLSPKPSPLQOEPSQSPSPSPICMMNALV 403
Db      319 CQCFRQKCLSGVMKVEVVRTDSLKGRRLSPKPSPLQOEPSQSPSPSPICMMNALV 369
QY      404 RALTDSTP--RDLDSRYCPTDQAAAG--TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTD 460
Db      370 RAHLDSPGSTAKLDYSKQELVLPHFGEKEDAGDVQOQFYDLLSGSLEVIKWAEEKIPGFAE 429
QY      461 LPKEDOTLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISKDF 520
Db      430 LSPGDQDLLESFALEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISKDF 489
QY      521 SINLOSINLDIQALACLSALSMTIRHGLKEPKRVEELCNKITSSLKDHQSKQALEPTE 580
Db      490 SRSLSGLVVDPAFACLSALVLTDRHGLQEPFRVEELQNRVIRASCLKEHVS-AVAGEPOP 548
QY      581 ----SKVLGALVELRKICTGLQRIFYKLEDLVSPSPSIIDKLFDTLPP 626
Db      549 ASCLSRLLGKLPRLTCTQGLQRIFYKLEDLVSPSPSIIDKLFDTLPP 598

RESULT 13
Q8N3V2
ID Q8N3V2 PRELIMINARY; PRT; 598 AA.
AC Q8N3V2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein DKFp451G125.
GN Name=DKFp451G125;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeinger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AL831844; CAD38550.2; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 598 AA; 64445 MW; C96154E06013AC8B CRC64;

Query Match 41.0%; Score 1368; DB 2; Length 598;
Best Local Similarity 48.5%; Pred. No. 1.2e-64;
Matches 315; Conservative 80; Mismatches 179; Indels 76; Gaps 21;

QY      1 MPCVQAQY---SPSPGGSSAAQYSSYETTIMNPDYTKLTMDLGSTEITATATSLPS 57
Db      1 MPCVQAQYTPAPSPGPDHDLA-----SDPLTPERIKPTMDLASPEAAPATLPS 52
QY      58 ISTFVEGYSNVVELKPCVYQMQQ---RPLKVEGRASRYHHHHHHHHHHHHHQQOQQ 114
Db      53 FTTFMDGVTGEFD---TFYQLPGVTVQPCSSASSASSSTSSSSATSPASAFKFDQVY 109
QY      115 PSIP-PASSPEPEVLPTSMYFKQPPS--TPTTAPFPQAGALWDEAL-PSAPGCIAPG 170

```

```

Db      110 GCYPGRSPGVDEALSSSGSDYDGSPCAPSPSTSPFQPPQLSPWDGSGHSSPSQTYEG 169
QY      171 PLLDPEMKAVPTVAGARFPLFHFKSPSPHPAPSPAGGHLGYDPTAAALSL-PLGAAA 229
Db      170 --LRATWQLPKASGPPQPAFFSPSPPTGSPS-----LAQSPKLFPQATH 216
QY      230 AAGSOAAALESHYPYGLPLAKRAAPLAPPLGLTPTSPASSLLG-----ESPSPSPSPSS 285
Db      217 QLQ-----EGESYSM-----PTAFPL-----APTSPHLESGILDTP-VTSTKARSG 258
QY      286 SSG--EGTCAVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNR 343
Db      259 APGSGEACVCGDNAAACHYGVRTCEGCKGFFKRTVQKNAYVCLANKNCVDRKRRNR 318
QY      344 CQCFRQKCLSGVMKVEVVRTDSLKGRRLSPKPSPLQOEPSQSPSPSPICMMNALV 403
Db      319 CQCFRQKCLSGVMKVEVVRTDSLKGRRLSPKPSPLQOEPSQSPSPSPICMMNALV 369
QY      404 RALTDSTP--RDLDSRYCPTDQAAAG--TDAEHVQOQFYNNLTASIDVSRSWAEKIPGFTD 460
Db      370 RAHLDSPGSTAKLDYSKQELVLPHFGEKEDAGDVQOQFYDLLSGSLEVIKWAEEKIPGFAE 429
QY      461 LPKEDOTLIESAFLEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISKDF 520
Db      430 LSPGDQDLLESFALEFVLRLSIRSNATAEKDFVFCNGLVLRHLOCLRGFGWLDISKDF 489
QY      521 SINLOSINLDIQALACLSALSMTIRHGLKEPKRVEELCNKITSSLKDHQSKQALEPTE 580
Db      490 SRSLSGLVVDPAFACLSALVLTDRHGLQEPFRVEELQNRVIRASCLKEHVA-AVAGEPOP 548
QY      581 ----SKVLGALVELRKICTGLQRIFYKLEDLVSPSPSIIDKLFDTLPP 626
Db      549 ASCLSRLLGKLPRLTCTQGLQRIFYKLEDLVSPSPSIIDKLFDTLPP 598

RESULT 14
NR41_HUMAN
ID NR41_HUMAN STANDARD; PRT; 598 AA.
AC P22736;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Orphan nuclear receptor NR4A1 (Orphan nuclear receptor HMR) (Early
DE response protein NAK1) (TR3 orphan nuclear receptor).
GN Name=NR4A1; Synonyms=GFRP1, HMR, NAK1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal skeletal muscle;
RX MEDLINE=91133413; PubMed=2283997;
RA Nakai A., Kartha S., Sakurai A., Toback F.G., Degroot L.J.;
RT "A human early response gene homologous to murine nur77 and rat NGFI-
RL B, and related to the nuclear receptor superfamily.";
RL Mol. Endocrinol. 4:1438-1443(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90173205; PubMed=2626032; DOI=10.1016/0022-4731(89)90114-3;
RA Chang C., Kokontis J., Liao S., Chang Y.;
RT "Isolation and characterization of human TR3 receptor: a member of
RL J. Steroid Biochem. 34:391-395(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Klausner R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wang J., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

Qy	1	MPCVQAQY-----SRSPPGSSVAAQYYSYSTEINWPDYTKLTMDLGSITEITATATSLPS	57
Db	1	MPCIOAQYGTAFSPGGRDHLA-----SDLPTEFIKPTMDLASPAAPAAFTALPS	52
Qy	58	ISTFVEGYSNNYELKPCSVYQM-Q--RPLIKVEBGRAPSYHHHHHHHHHHHHHHHHHHQHQ	114
Db	53	FSFTMDYGTGEFD---TFLVQLPCTVQPCSSASSASSSTSSSSATSPASAKFKDFQVY	109
Qy	115	PSIP-PASSPEDEVLPTSMYFKOSPPS--TPTTAPF-PPQAG-----	153
Db	110	GCYFPGPLSGPVDEALSSGSDYYGSPCSAPSPSTPSFQPPQLSPWDGSGFHGSPSQTYBG	169
Qy	154	-ALWDEALPAPGCIAPGLPDDPNKAVPTVAGARFLFHFKPSPPHPAPSPAGCHHLG	212
Db	170	LRAWTEQLPKASG-----PPQ-----PPAFFSFSPTGSPS-----	201
Qy	213	YDPTAAALSL-PLGAAAAAGSQAALLESHPYGLPLAKRAAPLAPPLGLTSPSTASSLL	271
Db	202	---LAGSPLKLFQATHQLG-----EGESYSM-----PTAPPGL-----APTSPHLE	241
Qy	272	G-----BSPSLSPSPSRSSSG--EGTCAVCGDNAAACQHYGVRTCEGCKGFFKRTVQKNK	325
Db	242	SGSILDTP-VTSTKARSGAPGSGSEGRCAVCGDNASCQHYGVRTCEGCKGFFKRTVQKNK	300
Qy	326	YVCLANKNCVPDKRRNRNCQYCRFKQKLSVGWKEVVRTDSLKGRRGRLPSKPKSPLOQE	385
Db	301	YICLANKDCVPDKRRNRNCQYCRFKQKLVAGWKEVVRTDSLKGRRGRLPSKPKQP----	356
Qy	386	PSQSPSPSPICMNNALVRALTOSTP--RDLDSRYCPTDQAAAG-TDAEHVQQFVNLIT	442
Db	357	-----PASPANLUTSLVRAHLDSGSPSTAKLDYSKFQELVLPHFGKEDAGDVQQFYDLLS	411
Qy	443	ASIDVSRSWAEKIPGFTDLPKEDQTLTIESAFLELFLVRLSIRNSTABDKFVFCNGLVLH	502
Db	412	GSLEIVRKWAEKIPGFAELSPADQDLLESFALELFLRLAYRSKPGSGKLIFCSGLVLH	471
Qy	503	RLOCLRGFGWLSIKDPSNLQSLNDIOALACLSALSMITERHGLKPKRVEBELCNKI	562
Db	472	RLOCARGFQGDWISILAFSRSLHLLVDVPFAFCLLSALVLITDRHGLQEPRRVEBELQNR	531
Qy	563	TSLSKHQHSKGQALEPTE-----SKVLGALVELRICTGLGQRIFYVLKLEDLVSPPSIIDK	618
Db	532	ASCLKEHVA-AVAGEPOPASCLSLRLGLKPELRLTCTGLQRIFYFKLEDLVPPPIIDK	590
Qy	619	FLDLDLTPF 626	
Db	591	IFNDLTPF 598	
RESULT 15			
Q6TBUB PRELIMINARY; PRT; 598 AA.			
ID	Q6IBUB		
AC	Q6IBUB		
DT	05-JUL-2004	(TrEMBLrel. 27, Created)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
DE	NR4A1	protein.	
GN	Name=NR4A1;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
ON	NCBI_TaxID=9606;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RP	Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;		
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: Nuclear (By similarity).		
CC	-1- SIMILARITY: Belongs to the nuclear hormone receptor family.		
DR	EMBL; CR456704; CAG32985.1; --		
DR	HSSP; P20393; IA6Y.		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0003707; F:steroid hormone receptor activity; IEA.		
DR	GO; GO:0003700; F:transcription factor activity; IEA.		

